

**Dossier zur Nutzenbewertung
gemäß § 35a SGB V**

Sarilumab (Kevzara®)

Sanofi-Aventis Deutschland GmbH

Modul 4 B

*Polymyalgia rheumatica (PMR) bei erwachsenen
Patienten, die auf Corticosteroide unzureichend
angesprochen haben oder bei denen ein Rezidiv
während des Ausschleichens der Corticosteroide auftritt*

Medizinischer Nutzen und
medizinischer Zusatznutzen,
Patientengruppen mit therapeutisch
bedeutsamem Zusatznutzen

Inhaltsverzeichnis

	Seite
Inhaltsverzeichnis	1
Tabellenverzeichnis	4
Abbildungsverzeichnis	9
Abkürzungsverzeichnis	11
4 Modul 4 – allgemeine Informationen	16
4.1 Zusammenfassung der Inhalte von Modul 4.....	17
4.2 Methodik.....	24
4.2.1 Fragestellung.....	24
4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung.....	26
4.2.3 Informationsbeschaffung.....	28
4.2.3.1 Studien des pharmazeutischen Unternehmers.....	28
4.2.3.2 Bibliografische Literaturrecherche.....	28
4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken.....	29
4.2.3.4 Suche auf der Internetseite des G-BA.....	31
4.2.3.5 Selektion relevanter Studien.....	32
4.2.4 Bewertung der Aussagekraft der Nachweise.....	32
4.2.5 Informationssynthese und -analyse.....	34
4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien.....	34
4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien.....	35
4.2.5.3 Meta-Analysen.....	50
4.2.5.4 Sensitivitätsanalysen.....	51
4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren.....	51
4.2.5.6 Indirekte Vergleiche.....	54
4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen.....	57
4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel.....	57
4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel.....	57
4.3.1.1.1 Studien des pharmazeutischen Unternehmers.....	57
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche.....	59
4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken.....	61
4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA.....	63
4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel.....	63
4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel.....	65
4.3.1.2.1 Studiendesign und Studienpopulationen.....	65
4.3.1.2.2 Verzerrungspotenzial auf Studienebene.....	74
4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien.....	75
4.3.1.3.1 Endpunkt – RCT.....	76
4.3.1.3.1.1 Endpunkt Gesamtmortalität - RCT.....	79

4.3.1.3.1.2	Endpunkt Anhaltende Remission zu Woche 52 - RCT	81
4.3.1.3.1.3	Endpunkt Kumulative Corticosteroid-Gesamtdosis über 52 Wochen - RCT	85
4.3.1.3.1.4	Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission - RCT	87
4.3.1.3.1.5	Endpunkt PMR-Aktivitätsscore (PMR-AS) - RCT	91
4.3.1.3.1.6	Patientenberichtete Endpunkte	99
4.3.1.3.1.6.1	Gesundheitszustand gemäß FACIT-F - RCT	99
4.3.1.3.1.6.2	Gesundheitszustand gemäß EQ-5D VAS - RCT	102
4.3.1.3.1.6.3	Körperlicher Funktionsstatus gemäß HAQ-DI - RCT	105
4.3.1.3.1.6.4	Gesundheitsbezogene Lebensqualität gemäß SF-36v2 - RCT	118
4.3.1.3.1.7	Endpunkt Nebenwirkungen - RCT	122
4.3.1.3.2	Subgruppenanalysen – RCT	131
4.3.1.3.2.1	Subgruppenanalysen für den Endpunkt Anhaltende Remission zu Woche 52 – RCT	136
4.3.1.3.2.2	Subgruppenanalysen für den Endpunkt Kumulative Corticosteroid-Gesamtdosis über 52 Wochen – RCT	140
4.3.1.3.2.3	Subgruppenanalysen für den Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission – RCT	140
4.3.1.3.2.4	Subgruppenanalysen für den Endpunkt PMR-AS – RCT	142
4.3.1.3.2.5	Subgruppenanalysen für den Endpunkt Veränderung des FACIT-F um $\geq 7,8$ Punkte – RCT	147
4.3.1.3.2.6	Subgruppenanalysen für den Endpunkt Veränderung der EQ- 5D VAS – RCT	148
4.3.1.3.2.7	Subgruppenanalysen für den Endpunkt Veränderung des HAQ- DI – RCT	148
4.3.1.3.2.8	Subgruppenanalysen für den Endpunkt Veränderung der SF- 36v2 – RCT	150
4.3.1.3.2.9	Subgruppenanalysen für den Endpunkt Nebenwirkungen – RCT	150
4.3.1.4	Liste der eingeschlossenen Studien - RCT	153
4.3.2	Weitere Unterlagen	153
4.3.2.1	Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien	153
4.3.2.1.1	Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche	153
4.3.2.1.2	Charakteristika der Studien für indirekte Vergleiche	154
4.3.2.1.3	Ergebnisse aus indirekten Vergleichen	154
4.3.2.1.3.1	<Endpunkt xxx> – indirekte Vergleiche aus RCT	154
4.3.2.1.3.2	Subgruppenanalysen – indirekte Vergleiche aus RCT	157
4.3.2.1.4	Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT	157
4.3.2.2	Nicht randomisierte vergleichende Studien	157
4.3.2.2.1	Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien	157
4.3.2.2.2	Charakteristika der nicht randomisierten vergleichenden Studien	158
4.3.2.2.3	Ergebnisse aus nicht randomisierten vergleichenden Studien	159
4.3.2.2.3.1	<Endpunkt xxx> – nicht randomisierte vergleichende Studien	159
4.3.2.2.3.2	Subgruppenanalysen – nicht randomisierte vergleichende Studien	160

4.3.2.2.4	Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien	160
4.3.2.3	Weitere Untersuchungen.....	160
4.3.2.3.1	Ergebnis der Informationsbeschaffung – weitere Untersuchungen	161
4.3.2.3.2	Charakteristika der weiteren Untersuchungen	161
4.3.2.3.3	Ergebnisse aus weiteren Untersuchungen	162
4.3.2.3.3.1	<Endpunkt xxx> – weitere Untersuchungen	162
4.3.2.3.3.2	Subgruppenanalysen – weitere Untersuchungen	163
4.3.2.3.4	Liste der eingeschlossenen Studien – weitere Untersuchungen.....	163
4.4	Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens.....	163
4.4.1	Beurteilung der Aussagekraft der Nachweise	163
4.4.2	Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß.....	165
4.4.3	Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht	171
4.5	Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte	172
4.5.1	Begründung für die Vorlage indirekter Vergleiche.....	172
4.5.2	Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen.....	172
4.5.3	Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen	172
4.5.4	Verwendung von Surrogatendpunkten	172
4.6	Referenzliste.....	174
Anhang 4-A : Suchstrategien – bibliografische Literaturrecherche		179
Anhang 4-B : Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken.....		183
Anhang 4-C : Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche).....		185
Anhang 4-D : Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken).....		186
Anhang 4-E : Methodik der eingeschlossenen Studien – RCT		187
Anhang 4-F : Bewertungsbögen zur Einschätzung von Verzerrungsaspekten		201
Anhang 4-G : Weitere Analysen		223

Tabellenverzeichnis

	Seite
Tabelle 4-1: Übersicht der Ein- und Ausschlusskriterien für RCT mit dem zu bewertenden Arzneimittel gegenüber der zVT	18
Tabelle 4-2: Ergebnisse, Aussagesicherheit und Ausmaß des Zusatznutzens von Sarilumab gegenüber der zVT auf Endpunktebene.....	20
Tabelle 4-3: Ergebnis, Aussagesicherheit und Ausmaß des Zusatznutzens von Sarilumab gegenüber der zVT für den Endpunkt Kumulative CS-Gesamtdosis	22
Tabelle 4-4: Übersicht der Ein- und Ausschlusskriterien für RCT mit dem zu bewertenden Arzneimittel gegenüber der zVT	27
Tabelle 4-5: Demografische Charakterisierung der Patientenpopulation zu Baseline.....	35
Tabelle 4-6: Krankheitsspezifische Charakterisierung der Patientenpopulation zu Baseline..	36
Tabelle 4-7: Patientenrelevante Endpunkte zum Nachweis des Zusatznutzens von Sarilumab im vorliegenden AWG.....	39
Tabelle 4-8: A priori definierte Subgruppenanalysen der Studie SAPHYR	52
Tabelle 4-9: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel	58
Tabelle 4-10: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel.....	58
Tabelle 4-11: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel.....	62
Tabelle 4-12: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel.....	63
Tabelle 4-13: Studienpool – RCT mit dem zu bewertenden Arzneimittel.....	64
Tabelle 4-14: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel	66
Tabelle 4-15: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel.....	67
Tabelle 4-16: Demografische Charakterisierung der Studienpopulation – RCT mit dem zu bewertenden Arzneimittel	67
Tabelle 4-17: Krankheitsspezifische Charakterisierung der Studienpopulation – RCT mit dem zu bewertenden Arzneimittel.....	69
Tabelle 4-18: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel.....	74
Tabelle 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel	75
Tabelle 4-20: Operationalisierung des Endpunkts Gesamtmortalität	79

Tabelle 4-21: Bewertung des Verzerrungspotenzials für den Endpunkt Gesamtmortalität in RCT mit dem zu bewertenden Arzneimittel	79
Tabelle 4-22: Ergebnisse für Gesamtmortalität aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	80
Tabelle 4-23: Operationalisierung des Endpunkts Anhaltende Remission zu Woche 52.....	81
Tabelle 4-24: Bewertung des Verzerrungspotenzials für den Endpunkt Anhaltende Remission zu Woche 52 in RCT mit dem zu bewertenden Arzneimittel	82
Tabelle 4-25: Ergebnisse für Anhaltende Remission zu Woche 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	83
Tabelle 4-26: Ergebnisse der Einzelkomponenten für Anhaltende Remission zu Woche 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	84
Tabelle 4-27: Operationalisierung des Endpunkts Kumulative CS-Gesamtdosis.....	85
Tabelle 4-28: Bewertung des Verzerrungspotenzials für den Endpunkt Kumulative CS-Gesamtdosis in RCT mit dem zu bewertenden Arzneimittel.....	85
Tabelle 4-29: Ergebnisse für Kumulative CS-Gesamtdosis aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	86
Tabelle 4-30: Operationalisierung des Endpunkts Zeit bis zum ersten PMR-Schub nach klinischer Remission	87
Tabelle 4-31: Bewertung des Verzerrungspotenzials für den Zeit bis zum ersten PMR-Schub nach klinischer Remission in RCT mit dem zu bewertenden Arzneimittel.....	88
Tabelle 4-32: Ergebnisse für Zeit bis zum ersten PMR-Schub nach klinischer Remission aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	89
Tabelle 4-33: Operationalisierung des Endpunkts PMR-AS	91
Tabelle 4-34: Bewertung des Verzerrungspotenzials für den Endpunkt PMR-AS in RCT mit dem zu bewertenden Arzneimittel	92
Tabelle 4-35: Veränderung des PMR-AS zu Woche 52 gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel.....	93
Tabelle 4-36: Veränderung der Einzelkomponenten des PMR-AS zu Woche 52 gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel.....	95
Tabelle 4-37: Rücklaufquoten der VAS – Schmerz, Patientenbeurteilung aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	97
Tabelle 4-38: Veränderung der Einzelkomponente VAS um $\geq -1,5$ Punkte gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)	98
Tabelle 4-39: Operationalisierung des Endpunkts Gesundheitszustand gemäß FACIT-F	99
Tabelle 4-40: Bewertung des Verzerrungspotenzials für den Endpunkt Gesundheitszustand gemäß FACIT-F in RCT mit dem zu bewertenden Arzneimittel.....	100
Tabelle 4-41: Rücklaufquoten des FACIT-F aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	100
Tabelle 4-42: Veränderung des FACIT-F um $\geq 7,8$ Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	101

Tabelle 4-43: Operationalisierung des Endpunkts Gesundheitszustand gemäß EQ-5D VAS	102
Tabelle 4-44: Bewertung des Verzerrungspotenzials für den Endpunkt Gesundheitszustand gemäß EQ-5D VAS in RCT mit dem zu bewertenden Arzneimittel	103
Tabelle 4-45: Rücklaufquoten der EQ-5D VAS aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	103
Tabelle 4-46: Veränderung der EQ-5D VAS um ≥ 15 Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	104
Tabelle 4-47: Operationalisierung des Endpunkts Körperlicher Funktionsstatus gemäß HAQ-DI.....	105
Tabelle 4-48: Bewertung des Verzerrungspotenzials für den Endpunkt Körperlicher Funktionsstatus gemäß HAQ-DI in RCT mit dem zu bewertenden Arzneimittel	106
Tabelle 4-49: Rücklaufquoten des HAQ-DI aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	107
Tabelle 4-50: Veränderung des Standard-Scores des HAQ-DI um $\geq -0,45$ Punkte und der globalen Krankheitsaktivität um ≥ -15 Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	107
Tabelle 4-51: Veränderung der Kategorien des HAQ-DI zu Woche 52 gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	108
Tabelle 4-52: Operationalisierung des Endpunkts Gesundheitsbezogene Lebensqualität gemäß SF-36v2	118
Tabelle 4-53: Bewertung des Verzerrungspotenzials für den Endpunkt Gesundheitsbezogene Lebensqualität gemäß SF-36v2 in RCT mit dem zu bewertenden Arzneimittel.....	119
Tabelle 4-54: Rücklaufquoten des SF-36v2 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	119
Tabelle 4-55: Veränderung der PCS und MCS des SF-36v2 um ≥ 10 Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	120
Tabelle 4-56: Operationalisierung des Endpunkts Nebenwirkungen.....	122
Tabelle 4-57: Bewertung des Verzerrungspotenzials für den Endpunkt Nebenwirkungen in RCT mit dem zu bewertenden Arzneimittel	123
Tabelle 4-58: Ergebnisse der sicherheitsrelevanten Endpunkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	123
Tabelle 4-59: Ergebnisse der UE gesamt nach SOC aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	125
Tabelle 4-60: Ergebnisse der UE gesamt nach PT aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	127
Tabelle 4-61: Ergebnisse der SUE nach SOC aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	128

Tabelle 4-62: Ergebnisse für UE, die zum Therapieabbruch führten, nach SOC und PT aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)	129
Tabelle 4-63: Ergebnisse für UE von besonderem Interesse mit PT aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)	130
Tabelle 4-64: Matrix der durchgeführten Subgruppenanalysen (Studie SAPHYR).....	132
Tabelle 4-65: Interaktionsterme (p-Werte) der Subgruppenanalysen je Endpunkt für Studie SAPHYR und Subgruppenmerkmale.....	133
Tabelle 4-66: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Remission spätestens zu Woche 12 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)	136
Tabelle 4-67: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Kein Krankheitsschub von Woche 12 bis 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	137
Tabelle 4-68: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Anhaltende CRP-Reduktion von Woche 12 bis 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	138
Tabelle 4-69: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Erfolgreiches Prednison-Ausschleichen von Woche 12 bis 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	139
Tabelle 4-70: Ergebnisse der Subgruppenanalysen für die Zeit bis zum ersten PMR-Schub nach klinischer Remission aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)	140
Tabelle 4-71: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Veränderung der VAS um $\geq -1,5$ Punkte für Krankheitsaktivität, Prüfarztbeurteilung, aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)	142
Tabelle 4-72: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Dauer der Morgensteifigkeit aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel.....	143
Tabelle 4-73: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Beweglichkeit der oberen Gliedmaßen aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	145
Tabelle 4-74: Ergebnisse der Subgruppenanalysen für die Veränderung des FACIT-F um $\geq 7,8$ Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)	147
Tabelle 4-75: Ergebnisse der Subgruppenanalysen für die Veränderung des Standard-Scores des HAQ-DI um $\geq -0,45$ Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	148
Tabelle 4-76: Ergebnisse der Subgruppenanalysen für die Veränderung der globalen Krankheitsaktivität des HAQ-DI um ≥ -15 Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	149
Tabelle 4-77: Ergebnisse der Subgruppenanalysen für UE gesamt aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)	150

Tabelle 4-78: Ergebnisse der Subgruppenanalysen für SUE aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation).....	152
Tabelle 4-79: Übersicht der eingeschlossenen Studie SAPHYR	153
Tabelle 4-80: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche	154
Tabelle 4-81: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden.....	155
Tabelle 4-82: Operationalisierung von <Endpunkt xxx>	155
Tabelle 4-83: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche	155
Tabelle 4-84: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche.....	156
Tabelle 4-85: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien	158
Tabelle 4-86: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien	159
Tabelle 4-87: Operationalisierung von <Endpunkt xxx>.....	159
Tabelle 4-88: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien	159
Tabelle 4-89: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen	162
Tabelle 4-90: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen.....	162
Tabelle 4-91: Ergebnisse, Aussagesicherheit und Ausmaß des Zusatznutzens von Sarilumab gegenüber der zVT auf Endpunktebene.....	166
Tabelle 4-92: Ergebnis, Aussagesicherheit und Ausmaß des Zusatznutzens von Sarilumab gegenüber der zVT für den Endpunkt Kumulative CS-Gesamtdosis	168
Tabelle 4-93: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens.....	171
Tabelle 4-94 (Anhang): Studiendesign und -methodik für Studie SAPHYR	188
Tabelle 4-95 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SAPHYR.....	202

Abbildungsverzeichnis

	Seite
Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel	61
Abbildung 4-2: Kaplan-Meier-Kurve für den Endpunkt „Zeit bis zum ersten PMR-Schub nach klinischer Remission“ aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)	90
Abbildung 4-3: Longitudinalplot für die Veränderung des PMR-AS gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	94
Abbildung 4-4: Longitudinalplot für die Veränderung des CRP gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel.....	96
Abbildung 4-5: Longitudinalplot für die Veränderung der Dauer der Morgensteifigkeit gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	96
Abbildung 4-6: Longitudinalplot für die Veränderung der Beweglichkeit der oberen Gliedmaßen gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel.....	97
Abbildung 4-7: Longitudinalplot für die Veränderung der Kategorie Ankleiden und Körperpflege des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	110
Abbildung 4-8: Longitudinalplot für die Veränderung der Kategorie Aufstehen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	111
Abbildung 4-9: Longitudinalplot für die Veränderung der Kategorie Essen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	112
Abbildung 4-10: Longitudinalplot für die Veränderung der Kategorie Gehen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	113
Abbildung 4-11: Longitudinalplot für die Veränderung der Kategorie Hygiene des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	114
Abbildung 4-12: Longitudinalplot für die Veränderung der Kategorie Erreichen von Gegenständen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	115
Abbildung 4-13: Longitudinalplot für die Veränderung der Kategorie Greifen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel	116
Abbildung 4-14: Longitudinalplot für die Veränderung der Kategorie Andere Tätigkeiten des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel.....	117
Abbildung 4-15: Kaplan-Meier-Kurve für den Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission, Subgruppenmerkmal „Altersgruppe (Jahre)“, aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population).....	141
Abbildung 4-16: Longitudinalplot für die Veränderung der Einzelkomponente Dauer der Morgensteifigkeit, Subgruppenmerkmal „Abstammung“, aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)	144

Abbildung 4-17: Longitudinalplot für die Veränderung der Einzelkomponente
Beweglichkeit der oberen Gliedmaßen, Subgruppenmerkmal „CRP zu Baseline (mg/l)“,
aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)..... 146

Abbildung 4-18: Flow Chart der Studie SAPHYR 200

Abkürzungsverzeichnis

Abkürzung	Bedeutung
ALT	Alanin-Aminotransferase
AMIce	Arzneimittelinformationssystem
AM-NutzenV	Arzneimittel-Nutzenbewertungsverordnung
ANC	Absolute Neutrophilenzahl (Absolute Neutrophil Count)
Anti-CCP	Anti-zyklisches citrulliniertes Peptid (Anti-Cyclic Citrullinated Peptide)
AWG	Anwendungsgebiet
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
BMI	Body Mass Index
CHMP	Ausschuss für Humanarzneimittel (Committee for Medicinal Products for Human Use)
CONSORT	Consolidated Standards of Reporting Trials
COVID-19	Corona-Virus Disease 2019
CRP	C-reaktives Protein
CS	Corticosteroide
CST	Corticosteroid-Ausschleichen (Corticosteroid Taper)
CTCAE	Common Terminology Criteria for Adverse Events
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
DMARD	Krankheitsmodifizierendes Antirheumatikum (Disease-Modifying Anti-Rheumatic Drug)
EG	Europäische Gemeinschaft
EMA	Europäische Arzneimittel-Agentur (European Medicines Agency)
EMBASE	Excerpta Medica dataBASE
EQ-5D	European Quality of Life 5 Dimensions
ESR	Erythrozyten-Sedimentationsrate
EU-CTR	European Union Clinical Trials Register
EUL/EULAR	European League Against Rheumatism
FACIT-F	Functional Assessment of Chronic Illness Therapy-Fatigue

Abkürzung	Bedeutung
FACIT-G	Functional Assessment of Chronic Illness Therapy-General
FPFV	Erster Patient erste Visite (First Patient First Visit)
G-BA	Gemeinsamer Bundesausschuss
gp130	Glykoprotein 130
GTI	Glukokortikoid-Toxizitätsindex
HAQ-DI	Health Assessment Questionnaire-Disability Index
HbA1c	Glykiertes Hämoglobin
HBcAb	Hepatitis-B-Core-Antikörper (Hepatitis B Core Antibody)
HBsAg	Hepatitis-B-Oberflächenantigen (Hepatitis B Surface Antigen)
HBV DNA	Hepatitis B Virus Deoxyribunucleic Acid
HIV	Humanes Immundefizienz-Virus
HR	Hazard Ratio
ICTRP	International Clinical Trials Registry Platform
IgG1	Immunglobulin G1
IL-6	Interleukin-6
IL-6R	IL-6-Rezeptor
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
IRT	Interaktive Dialogtechnologie (Interactive Response Technology)
ITT	Intention-To-Treat
k.A.	Keine Angabe
KI	Konfidenzintervall
LPLV	Letzter Patient letzte Visite (Last Participant Last Visit)
LSMean	Kleinste-Quadrate-Mittelwert (Least Squares Mean)
LS-MW	Kleinste-Quadrate-Mittelwert (Least Squares Mean)
LS-MWD	Kleinste-Quadrate-Mittelwertdifferenz (Least Squares-Mean Difference)
Max	Maximum

Abkürzung	Bedeutung
MCS	Summenskala der psychischen Gesundheit (Mental Health Component Summary Score)
MedDRA	Medizinisches Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung (Medical Dictionary for Regulatory Activities)
MEDLINE	Medical Literature Analysis and Retrieval System Online
Min	Minimum
MMRM	Mixed Model for Repeated Measures
MTC	Mixed Treatment Comparison
MTX	Methotrexat
MW	Mittelwert
N	Anzahl der Patienten
n	Anzahl der Patienten mit Ereignis / Ausprägung / Ergebnis
NA	Nicht anwendbar
NCT	National Clinical Trial
NE	Nicht berechenbar (Not Evaluable)
NI	Nicht interpretierbar
OR	Odds Ratio
PCS	Summenskala der körperlichen Gesundheit (Physical Health Component Summary Score)
PICOS	Population, Intervention, Comparison, Outcome, Study
PMR	Polymyalgia rheumatica
PMR-AS	Polymyalgia rheumatica-Aktivitätsscore
PRO	Patientenberichtete Endpunkte (Patient-Reported Outcomes)
PT	Bevorzugter Begriff nach MedDRA (Preferred Term)
Q	Quartil
q2w	Alle zwei Wochen
RCT	Randomisierte, kontrollierte Studie (Randomized Controlled Trial)
RD	Risikodifferenz
RF	Rheumafaktor

Abkürzung	Bedeutung
RR	Relatives Risiko
SAP	Statistischer Analyseplan
s.c.	Subkutan (Subcutaneous)
SD	Standardabweichung (Standard Deviation)
SE	Standardfehler (Standard Error)
SF-36v2	36-Item Short Form Health Survey Version 2
SGB	Sozialgesetzbuch
SMD	Standardisierte Mittelwertdifferenz
SMQs	Standardised MedDRA Queries
SOC	Systemorganklasse nach MedDRA (System Organ Class)
STE	Surrogate Threshold Effects
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SUE	Schwerwiegendes UE
TB	Tuberkulose
TREND	Transparent Reporting of Evaluations with Non-Randomized Design
UE	Unerwünschtes Ereignis
ULN	Obere Normgrenze (Upper Limit of Normal)
USA	Vereinigte Staaten von Amerika (United States of America)
VAS	Visuelle Analogskala
VerfO	Verfahrensordnung
WHO	Weltgesundheitsorganisation (World Health Organization)
zVT	Zweckmäßige Vergleichstherapie

Zur besseren Lesbarkeit wird in diesem Dossier auf die geschlechtsspezifische Darstellung verzichtet und das generische Maskulinum verwendet. Sämtliche Personenbezeichnungen gelten gleichermaßen für alle Geschlechter.

4 Modul 4 – allgemeine Informationen

Modul 4 enthält folgende Angaben:

- Zusammenfassung (Abschnitt 4.1)
- Angaben zur Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens (Abschnitt 4.2)
- Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen (Abschnitt 4.3)
- eine abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens, einschließlich der Angabe von Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht (Abschnitt 4.4)
- ergänzende Informationen zur Begründung der vorgelegten Unterlagen (Abschnitt 4.5)

Für jedes zu bewertende Anwendungsgebiet ist eine separate Version des vorliegenden Dokuments zu erstellen. Die Kodierung der Anwendungsgebiete ist in Modul 2 hinterlegt. Sie ist je Anwendungsgebiet einheitlich für die Module 3, 4 und 5 zu verwenden.

Im Dokument verwendete Abkürzungen sind in das Abkürzungsverzeichnis aufzunehmen. Sofern Sie für Ihre Ausführungen Tabellen und Abbildungen verwenden, sind diese im Tabellen- bzw. Abbildungsverzeichnis aufzuführen.

4.1 Zusammenfassung der Inhalte von Modul 4

Stellen Sie eine strukturierte Zusammenfassung der Inhalte von Modul 4 zur Verfügung.

Fragestellung

In dem vorliegenden Anwendungsgebiet (AWG) ist Sarilumab gemäß Fachinformation zur Behandlung der Polymyalgia rheumatica (PMR) bei erwachsenen Patienten angezeigt, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt.

Ziel des Dossiers ist die Bewertung des medizinischen Nutzens und Zusatznutzens von Sarilumab im vorliegenden AWG nach dem anerkannten Stand der medizinischen Erkenntnisse hinsichtlich der patientenrelevanten Endpunkte Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen. Die einzelnen Endpunkte sind in Abschnitt 4.2.5.2 detailliert beschrieben. Die Bewertung erfolgt gegenüber Placebo in Kombination mit Prednison. Aufgrund der Einstufung systemischer Corticosteroide (CS) als Option der zweckmäßigen Vergleichstherapie (zVT) durch den Gemeinsamen Bundesausschuss (G-BA) ist ein direkter Vergleich zur zVT möglich. Prednison wurde während der Behandlungsphase im Kontrollarm leitlinienkonform ausgeschlichen.

Datenquellen

Zur Identifizierung relevanter Datenquellen wurde eine bibliografische Literaturrecherche nach randomisierten, kontrollierten Studien (Randomized Controlled Trial, RCT) mit Sarilumab im vorliegenden AWG in den Datenbanken EMBASE, MEDLINE und dem Cochrane Central Register of Controlled Trials durchgeführt. Darüber hinaus wurden folgende Studienregister und Studienergebnisdatenbanken durchsucht: Clinicaltrials.gov, EU Clinical Trials Register (EU-CTR), International Clinical Trials Registry Platform (ICTRP) der Weltgesundheitsorganisation (WHO) und Arzneimittelinformationssystem (AMIce). Ergänzend wurde eine Suche auf der Internetseite des G-BA durchgeführt.

In der Recherche wurde die klinische Zulassungsstudie SAPHYR (EFC15160) im vorliegenden AWG von Sarilumab identifiziert, die zum Nachweis des Nutzens und Zusatznutzens herangezogen wird. Die Studie entspricht nach der in der Verfahrensordnung des G-BA dargestellten Evidenzhierarchie der Evidenzstufe Ib. Es handelt sich um eine abgeschlossene, multizentrische, randomisierte, doppelblinde, placebokontrollierte, parallele Phase-III-Studie, die Sarilumab in Kombination mit Prednison gegenüber Placebo in Kombination mit Prednison an Patienten im Alter ab 50 Jahren mit aktiver PMR untersucht. Als Datenquellen werden die Studienunterlagen sowie dazugehörige Zusatzauswertungen zu den Endpunkten und Subgruppenanalysen herangezogen. Die vollständigen Studienunterlagen sind in Modul 5 abgelegt.

Ein-/Ausschlusskriterien für Studien

Zur Selektion von relevanten Studien für die Nutzenbewertung wurden Ein- und Ausschlusskriterien in den Kategorien Patientenpopulation, Intervention, Vergleichstherapie, Endpunkte, Studientyp und -dauer, Sprache, Publikationstyp sowie Studienstatus definiert

(Tabelle 4-1). Studien, die ein oder mehrere Einschlusskriterien nicht erfüllten, wurden ausgeschlossen.

Tabelle 4-1: Übersicht der Ein- und Ausschlusskriterien für RCT mit dem zu bewertenden Arzneimittel gegenüber der zVT

Kategorie	Einschlusskriterien (E)		Ausschlusskriterien (A)	
Patientenpopulation	E1	Erwachsene Patienten mit PMR, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt	A1	Andere Patientenpopulation
Intervention	E2	Sarilumab (s.c. Injektion): 200 mg q2w	A2	Andere Dosierung oder Formulierung von Sarilumab oder andere Medikation
Vergleichstherapie	E3	Therapie nach ärztlicher Maßgabe unter Berücksichtigung von systemischen Glukokortikoiden und der Kombination von Glukokortikoiden mit Methotrexat	A3	Abweichende Vergleichsintervention
Endpunkte	E4	Mindestens ein patientenrelevanter Endpunkt aus den Kategorien <ul style="list-style-type: none"> • Mortalität • Morbidität • Gesundheitsbezogene Lebensqualität • Nebenwirkungen 	A4	Ungeeignete Endpunkte oder keine verwertbaren Ergebnisse zu patientenrelevanten Endpunkten
Studientyp	E5	RCT	A5	Andere Studien als RCT, tierexperimentelle Studien, in vitro-Studien
Studiendauer	E6	≥ 24 Wochen	A6	< 24 Wochen
Sprache	E7	Englisch oder Deutsch	A7	Andere als Englisch oder Deutsch
Publikationstyp	E8	Vollpublikation oder andere ausführliche Ergebnisberichte, z. B. Studienbericht, laufende Studie mit verfügbaren Zwischenergebnissen	A8	Mehrfachpublikationen ohne relevante Zusatzinformationen, Konferenz-Abstracts oder Poster
Studienstatus	E9	Studien, für die Ergebnisse vorliegen	A9	Noch nicht begonnene, laufende oder abgeschlossene Studien ohne publizierte Ergebnisse
PMR: Polymyalgia rheumatica; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); s.c.: Subkutan (Subcutaneous); zVT: Zweckmäßige Vergleichstherapie.				

Methoden zur Bewertung der Aussagekraft der Nachweise und zur Synthese von Ergebnissen

Die Bewertung der Aussagekraft der Nachweise erfolgte auf Studienebene und separat für jeden patientenrelevanten Endpunkt auf Basis der klinischen Studie SAPHYR. Die Darstellung der Studie gemäß Consolidated Standards of Reporting Trials (CONSORT) befindet sich in Anhang 4-E.

Die Dokumentation zur Einschätzung des Verzerrungspotenzials erfolgte unter Verwendung von Anhang 4-F. Das Verzerrungspotenzial sowohl auf Studienebene (A) als auch auf Endpunktebene (B) wurde nach den Vorgaben der Modulvorlage für die endpunktübergreifenden bzw. endpunktspezifischen Aspekte bewertet und als „niedrig“ oder „hoch“ eingestuft.

Ausgehend von der Beurteilung des Verzerrungspotenzials auf Studien- und Endpunktebene erfolgte zunächst die Graduierung des Ausmaßes der Ergebnissicherheit auf Studien- und Endpunktebene und im Anschluss daran die Klassifizierung der Wahrscheinlichkeit des Zusatznutzens für jeden patientenrelevanten Endpunkt in „Beleg“, „Hinweis“, „Anhaltspunkt“ oder „kein nachgewiesener Zusatznutzen“ gemäß den Angaben im Methodenpapier des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG).

Ergebnisse zum medizinischen Nutzen und medizinischen Zusatznutzen

Die Bewertung des medizinischen Nutzens und Zusatznutzens erfolgt auf Grundlage der Zulassungsstudie SAPHYR unter Anwendung der Kriterien der Arzneimittel-Nutzenbewertung (AM-NutzenV) für die Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen.

Als zVT bestimmte der G-BA „eine Therapie nach ärztlicher Maßgabe unter Berücksichtigung von systemischen Glukokortikoiden und der Kombination von Glukokortikoiden mit Methotrexat“.

In der Studie SAPHYR erhielten die Patienten zusätzlich zu Sarilumab bzw. Placebo ebenfalls Prednison, das im Sarilumab-Arm über 14 Wochen und im Kontrollarm leitlinienkonform über 52 Wochen ausgeschlichen wurde. Die Gabe von kommerziellen CS als Notfallmedikation war möglich. Die zVT wird in der Studie als adäquat umgesetzt betrachtet. Es wurden ebenso Patienten eingeschlossen, die begleitend zu Prednison nach ärztlicher Maßgabe Methotrexat (MTX) einnahmen. Das Verzerrungspotenzial auf Studienebene ist insgesamt als niedrig einzustufen. Mit der Studie SAPHYR steht somit eine direkt vergleichende hochwertige RCT mit der vom G-BA benannten zVT für die Nutzenbewertung von Sarilumab im AWG zur Verfügung.

Nachfolgend wird der Zusatznutzen anhand des Vorteils von Sarilumab in den patientenrelevanten Kategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen dargestellt (Tabelle 4-2).

Tabelle 4-2: Ergebnisse, Aussagesicherheit und Ausmaß des Zusatznutzens von Sarilumab gegenüber der zVT auf Endpunktebene

Nutzendimension/ Endpunkt	Sarilumab+CST n/N (%)	Placebo+CST n/N (%)	Sarilumab+CST vs. Placebo+CST	Aussagesicherheit und Ausmaß des Zusatznutzens
			Effektschätzer [95 %-KI]; p-Wert	
Gesamtmortalität				
Keine Todesfälle in der Studie SAPHYR				Nicht zur Ableitung des Zusatznutzens herangezogen
Morbidität				
Remission				
Anhaltende Remission zu Woche 52	17/60 (28,3)	6/58 (10,3)	RR 2,74 [1,15; 6,52]; 0,0227	Hinweis auf einen beträchtlichen Zusatznutzen
Einzelkomponenten der Remission				
Remission spätestens zu Woche 12	28/60 (46,7)	22/58 (37,9)	RR 1,23 [0,80; 1,89]; 0,3454	Zusatznutzen nicht belegt
Kein Krankheitsschub von Woche 12 bis 52	33/60 (55,0)	19/58 (32,8)	RR 1,68 [1,08; 2,60]; 0,0206	Hinweis auf einen geringen Zusatznutzen
Anhaltende CRP- Reduktion von Woche 12 bis 52	40/60 (66,7)	26/58 (44,8)	RR 1,49 [1,06; 2,09]; 0,0223	Nicht zur Ableitung des Zusatznutzens herangezogen
Erfolgreiches Prednison- Ausschleichen von Woche 12 bis 52	30/60 (50,0)	14/58 (24,1)	RR 2,07 [1,22; 3,51]; 0,0068	Hinweis auf einen beträchtlichen Zusatznutzen
PMR-Schub				
Zeit bis zum ersten PMR- Schub nach klinischer Remission	29/60 (48,3)	45/58 (77,6)	HR 0,56 [0,351; 0,895]; 0,0153	Hinweis auf einen beträchtlichen Zusatznutzen
PMR-AS				
Veränderung des PMR- AS zu Woche 52	31/60 (51,7)	28/58 (48,3)	LS-MWD -5,30 [-8,01; -2,60]; 0,0002 Hedges' g -1,01 [-1,52; -0,49]	Anhaltspunkt für einen beträchtlichen Zusatznutzen
Einzelkomponenten des PMR-AS				
Veränderung der VAS – Schmerz, Patientenbeurteilung (0 bis 10) um $\geq -1,5$ Punkte	20/60 (33,3)	20/58 (34,5)	RR 0,97 [0,58; 1,61]; 0,8962	Zusatznutzen nicht belegt
Veränderung der VAS – Krankheitsaktivität,	30/60 (50,0)	21/58 (36,2)	RR 1,38 [0,90; 2,12]; 0,1408	Zusatznutzen nicht belegt

Nutzendimension/ Endpunkt	Sarilumab+CST n/N (%)	Placebo+CST n/N (%)	Sarilumab+CST vs. Placebo+CST	Aussagesicherheit und Ausmaß des Zusatznutzens
			Effektschätzer [95 %-KI]; p-Wert	
Prüfarztbeurteilung (0 bis 10) um $\geq -1,5$ Punkte				
Veränderung des CRP (mg/dl) zu Woche 52	43/60 (71,7)	35/58 (60,3)	LS-MWD -7,10 [-11,91; -2,29]; 0,0042 Hedges' g -0,66 [-1,10; -0,21]	Nicht zur Ableitung des Zusatznutzens herangezogen
Veränderung der Dauer der Morgensteifigkeit (min x 0,1) zu Woche 52	48/60 (80,0)	46/58 (79,3)	LS-MWD -22,43 [-39,09; -5,77]; 0,0088 Hedges' g -0,55 [-0,96; -0,14]	Zusatznutzen nicht belegt
Veränderung der Beweglichkeit der oberen Gliedermaßen (EUL [3-0]) zu Woche 52	48/60 (80,0)	46/58 (79,3)	LS-MWD -0,24 [-0,40; -0,08]; 0,0044 Hedges' g -0,60 [-1,00; -0,19]	Zusatznutzen nicht belegt
Gesundheitszustand				
Veränderung des FACIT- F um $\geq 7,8$ Punkte	24/60 (40,0)	17/58 (29,3)	RR 1,36 [0,82; 2,28]; 0,2331	Zusatznutzen nicht belegt
Veränderung der EQ- 5D VAS um ≥ 15 Punkte	16/60 (26,7)	9/58 (15,5)	RR 1,72 [0,82; 3,60]; 0,1517	Zusatznutzen nicht belegt
Körperlicher Funktionsstatus				
Veränderung des HAQ- DI um $\geq -0,45$ Punkte	19/60 (31,7)	10/58 (17,2)	RR 1,84 [0,93; 3,63]; 0,0808	Zusatznutzen nicht belegt
Gesundheitsbezogene Lebensqualität				
Veränderung des SF-36v2				
Veränderung der PCS um ≥ 10 Punkte	14/60 (23,3)	10/58 (17,2)	RR 1,35 [0,65; 2,82]; 0,4194	Zusatznutzen nicht belegt
Veränderung der MCS um ≥ 10 Punkte	11/60 (18,3)	5/58 (8,6)	RR 2,13 [0,78; 5,80]; 0,1408	Zusatznutzen nicht belegt
Nebenwirkungen				
UE gesamt	56/59 (94,9)	49/58 (84,5)	RR 1,1 [0,99; 1,27]; 0,0711	Nicht zur Ableitung des Zusatznutzens herangezogen
SUE	8/59 (13,6)	12/58 (20,7)	RR 0,66 [0,29; 1,50]; 0,3164	Zusatznutzen nicht belegt

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Nutzendimension/ Endpunkt	Sarilumab+CST n/N (%)	Placebo+CST n/N (%)	Sarilumab+CST vs. Placebo+CST	Aussagesicherheit und Ausmaß des Zusatznutzens
			Effektschätzer [95 %-KI]; p-Wert	
Schwere UE	5/59 (8,5)	7/58 (12,1)	RR 0,70 [0,23; 2,11]; 0,5289	Zusatznutzen nicht belegt
UE, die zum Therapieabbruch führten	7/59 (11,9)	4/58 (6,9)	RR 1,72 [0,53; 5,63]; 0,3701	Zusatznutzen nicht belegt
UE von besonderem Interesse	7/59 (11,9)	5/58 (8,6)	RR 1,38 [0,46; 4,14]; 0,5695	Zusatznutzen nicht belegt

CRP: C-reaktives Protein; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); EQ-5D: European Quality of Life 5 Dimensions; EULAR: European League Against Rheumatism (EULAR); FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI: Health Assessment Questionnaire Disability Index; HR: Hazard Ratio; KI: Konfidenzintervall; LS-MWD: Kleinste-Quadrat-Mittelwertdifferenz (Least Square-Mean Value Difference); MCS: Summenskala der psychischen Gesundheit (Mental Health Component Summary Score); N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; PCS: Summenskala der körperlichen Gesundheit (Physical Health Component Summary Score); PMR: Polymyalgia rheumatica; PMR-AS: PMR-Aktivitätsscore; RR: Relatives Risiko; SF-36v2: 36-Item Short Form Health Survey, Version 2; SUE: Schwerwiegendes UE; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala; zVT: Zweckmäßige Vergleichstherapie.

Tabelle 4-3: Ergebnis, Aussagesicherheit und Ausmaß des Zusatznutzens von Sarilumab gegenüber der zVT für den Endpunkt Kumulative CS-Gesamtdosis

Nutzendimension/ Endpunkt	Sarilumab+CST Median (mg) [95 %-KI]	Placebo+CST Median (mg) [95 %-KI]	Sarilumab+CST vs. Placebo+CST	Aussagesicherheit und Ausmaß des Zusatznutzens
			p-Wert	
Morbidität				
Kumulative CS-Gesamtdosis				
Differenz zwischen tatsächlicher und erwarteter Gesamtdosis	0,0 [0,0; 0,0]	199,5 [0,0; 496,0]	0,0189	Nicht zur Ableitung des Zusatznutzens herangezogen

CS: Corticosteroids; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall;
zVT: Zweckmäßige Vergleichstherapie.

Die Ableitung des Zusatznutzens erfolgt für alle Patienten auf Basis der Gesamtpopulation (Hauptanalyse). Subgruppeneffekte werden gemäß den Vorgaben der Verfahrensordnung des G-BA in Abschnitt 4.3.1.3.2 dargestellt. Es konnten jedoch keine endpunktübergreifenden Effektmodifikationen festgestellt werden. Bei den vereinzelt beobachteten Subgruppeneffekten kann nicht abschließend bestimmt werden, ob es sich um eine tatsächliche Effektmodifikation oder um ein falsch-positives Ergebnis des Interaktionstests handelt. Somit bleiben die Subgruppenanalysen zur Ableitung des Zusatznutzens unberücksichtigt.

Schlussfolgerungen zum Zusatznutzen und zum therapeutisch bedeutsamen Zusatznutzen

Zusammenfassend zeigen sich statistisch signifikante Vorteile bei den Endpunkten „anhaltende Remission zu Woche 52“, „Zeit bis zum ersten PMR-Schub nach klinischer Remission“ und „Veränderung des PMR-Aktivitätsscore (PMR-AS) zu Woche 52“. Der signifikante Vorteil beim Endpunkt „anhaltende Remission zu Woche 52“ wird weitestgehend auch in der Analyse der Einzelkomponenten bestätigt. In der Gesamtschau ergibt sich demnach unter Berücksichtigung dieser Vorteile ein **Hinweis auf einen beträchtlichen Zusatznutzen** für Sarilumab gegenüber der zVT in der Behandlung erwachsener Patienten mit PMR, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt.

4.2 Methodik

Abschnitt 4.2 soll die Methodik der im Dossier präsentierten Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens beschreiben. Der Abschnitt enthält Hilfestellungen für die Darstellung der Methodik sowie einige Vorgaben, die aus den internationalen Standards der evidenzbasierten Medizin abgeleitet sind. Eine Abweichung von diesen methodischen Vorgaben ist möglich, bedarf aber einer Begründung.

4.2.1 Fragestellung

Nach den internationalen Standards der evidenzbasierten Medizin soll eine Bewertung unter einer definierten Fragestellung vorgenommen werden, die mindestens folgende Komponenten enthält:

- Patientenpopulation
- Intervention
- Vergleichstherapie
- Endpunkte
- Studientypen

Unter Endpunkte sind dabei alle für die frühe Nutzenbewertung relevanten Endpunkte anzugeben (d. h. nicht nur solche, die ggf. in den relevanten Studien untersucht wurden).

Die Benennung der Vergleichstherapie in Modul 4 muss zur Auswahl der zweckmäßigen Vergleichstherapie im zugehörigen Modul 3 konsistent sein.

Geben Sie die Fragestellung der vorliegenden Aufarbeitung von Unterlagen zur Untersuchung des medizinischen Nutzens und des medizinischen Zusatznutzens des zu bewertenden Arzneimittels an. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Sarilumab ist angezeigt für die Behandlung der Polymyalgia rheumatica (PMR) bei erwachsenen Patienten, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt .

PMR ist die zweithäufigste entzündlich-rheumatische immunvermittelte Erkrankung bei Personen, die älter sind als 50 Jahre, und äußert sich typischerweise durch Morgensteifigkeit und akute beidseitige Schmerzen vor allem in den Schultern, aber bei mehr als der Hälfte der Patienten auch im Nacken- und/oder im Beckengürtelbereich. Die Schmerzen und die Morgensteifigkeit führen meist zu einer bedeutenden Beeinträchtigung der Mobilität der Betroffenen und verhindern selbst Routinebewegungen wie das Aufstehen aus dem Sitzen oder Liegen, das Duschen, das Ankleiden oder das Autofahren. Aufgrund der beträchtlichen körperlichen Beeinträchtigung stellt PMR eine erhebliche Belastung für die Betroffenen dar und wirkt sich auch psychisch sehr negativ auf die Betroffenen aus (Acharya 2024; Buttgerit 2018; Lundberg 2022).

Ziel des vorliegenden Dossiers ist die Bewertung des medizinischen Nutzens und Zusatznutzens von Sarilumab zur Behandlung der PMR bei erwachsenen Patienten, die auf

Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt.

Die Bewertung erfolgt gegenüber der vom G-BA festgelegten zVT anhand patientenrelevanter Endpunkte der Nutzenkategorien Mortalität, Morbidität, Lebensqualität und Nebenwirkungen. Basis der Nutzenbewertung sind RCT.

Patientenpopulation

Sarilumab ist angezeigt für die Behandlung der PMR bei erwachsenen Patienten, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt (Sanofi 2024a).

Intervention

Sarilumab (Kevzara[®]) ist ein humaner monoklonaler Antikörper vom Typ Immunglobulin G1 (IgG1), der mit hoher Affinität an die α -Untereinheit sowohl löslicher als auch membrangebundener Interleukin-6-Rezeptoren (IL-6R α) bindet. Dadurch wird die Bindung von Interleukin-6 (IL-6) an seinen Rezeptor blockiert und die nachfolgende, über das Glykoprotein 130 (gp130) intrazellulär ausgelöste, Signalkaskade gehemmt. Entsprechend der Zulassung erfolgt die Behandlung mit Sarilumab mittels subkutaner Injektion von 200 mg alle zwei Wochen (q2w). Sarilumab wird in Kombination mit Corticosteroiden gegeben, die über einen durch den Arzt definierten Zeitraum ausgeschlichen werden (Sanofi 2024a).

Vergleichstherapie

Als zVT im vorliegenden AWG von Sarilumab bestimmte der G-BA (G-BA 2024):

- Eine Therapie nach ärztlicher Maßgabe unter Berücksichtigung von systemischen Glukokortikoiden und der Kombination von Glukokortikoiden mit Methotrexat

In Deutschland sind die Corticosteroide Prednison, Prednisolon, Methylprednisolon und Triamcinolon zur Behandlung von Patienten mit PMR grundsätzlich zugelassen (acis 2024a, 2024b; Dermapharm 2022; Galen 2024). Methotrexat besitzt keine Zulassung im vorliegenden AWG, dennoch sollte eine zusätzliche Gabe für Patienten mit unzureichendem Ansprechen auf Corticosteroide bzw. hohem Risiko für Rezidive und/oder langer Therapiedauer sowie Corticosteroid-induzierten Nebenwirkungen frühzeitig in Betracht gezogen werden (Buttgereit 2018; Dejaco 2015).

Endpunkte

Zur Bewertung des medizinischen Nutzens und Zusatznutzens von Sarilumab im vorliegenden AWG werden patientenrelevante Endpunkte aus den Nutzenkategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen entsprechend der AM-NutzenV dargestellt (BMJ 2023).

Die im Rahmen des Dossiers betrachteten patientenrelevanten Endpunkte aus den vorstehend genannten Nutzenkategorien werden in Abschnitt 4.2.5.2 detailliert dargestellt.

Studientypen

Die Bewertung des medizinischen Nutzens und Zusatznutzens von Sarilumab im vorliegenden AWG erfolgt anhand der pivotalen RCT der Phase-III-Studie SAPHYR (EFC15160).

Die Empfehlungen des Ausschusses für Humanarzneimittel (Committee for Medicinal Products for Human Use, CHMP) zur Zulassung von Sarilumab im vorliegenden AWG basieren auf den Ergebnissen der pivotalen klinischen Phase-III-Studie SAPHYR (EMA 2024).

4.2.2 Kriterien für den Einschluss von Studien in die Nutzenbewertung

Die Untersuchung der in Abschnitt 4.2.1 benannten Fragestellung soll auf Basis von klinischen Studien vorgenommen werden. Für die systematische Auswahl von Studien für diese Untersuchung sollen Ein- und Ausschlusskriterien für die Studien definiert werden. Dabei ist zu beachten, dass eine Studie nicht allein deshalb ausgeschlossen werden soll, weil keine in einer Fachzeitschrift veröffentlichte Vollpublikation vorliegt. Eine Bewertung der Studie kann beispielsweise auch auf Basis eines ausführlichen Ergebnisberichts aus einem Studienregister/ einer Studienergebnisdatenbank erfolgen, während ein Kongressabstrakt allein in der Regel nicht für eine Studienbewertung ausreicht.

Benennen Sie die Ein- und Ausschlusskriterien für Studien zum medizinischen Nutzen und Zusatznutzen. Machen Sie dabei mindestens Aussagen zur Patientenpopulation, zur Intervention, zur Vergleichstherapie, zu den Endpunkten, zum Studientyp und zur Studiendauer und begründen Sie diese. Stellen Sie die Ein- und Ausschlusskriterien zusammenfassend in einer tabellarischen Übersicht dar. Erstellen Sie dabei für unterschiedliche Themen der Recherche (z. B. unterschiedliche Fragestellungen) jeweils eine separate Übersicht.

Es wurden Ein- und Ausschlusskriterien definiert, um diejenigen Studien zu identifizieren, die zur Beantwortung der in Abschnitt 4.2.1 dargelegten Fragestellung geeignet sind. Die Selektionskriterien berücksichtigen die Kategorien Patientenpopulation, Intervention, Vergleichstherapie, Endpunkte, Studientyp und -dauer, Sprache, Publikationstyp sowie Studienstatus (Tabelle 4-4).

Die **Patientenpopulation** wird durch das zugelassene AWG definiert und umfasst erwachsene Patienten mit PMR, die auf Corticosteroide unzureichend angesprochen haben oder die bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt (Sanofi 2024a). Es wurden Studien eingeschlossen, in denen das zu bewertende Arzneimittel Sarilumab im vorliegenden AWG in der zugelassenen Art der Anwendung (**Intervention**) als subkutane Injektion von 200 mg q2w geprüft wurde. Als zweckmäßige **Vergleichstherapie** bestimmte der G-BA eine „Therapie nach ärztlicher Maßgabe unter Berücksichtigung von systemischen Glukokortikoiden und der Kombination von Glukokortikoiden mit Methotrexat“ (G-BA 2024). Die Bewertung des medizinischen Nutzens und Zusatznutzens von Sarilumab im vorliegenden AWG basiert auf patientenrelevanten **Endpunkten** aus den Kategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen. Es musste wenigstens ein patientenrelevanter Endpunkt aus einer der 4 Kategorien gegeben sein, um eine Bewertung zu

ermöglichen. Hinsichtlich **Studientyp**, **Publikationstyp** und **Sprache** wurden abgeschlossene oder laufende RCT eingeschlossen, zu denen Vollpublikationen, publizierte Zwischenergebnisse oder andere ausführliche Ergebnisberichte in deutscher oder englischer Sprache verfügbar sind. Nicht berücksichtigt wurden Studien, zu denen lediglich Berichte in Form von Abstracts, Kongressberichten, Postern oder in ähnlicher, unvollständiger Form vorliegen. Die **Studiendauer** wurde unter Berücksichtigung der Angaben des G-BA im Beratungsgespräch auf mindestens 24 Wochen festgelegt (G-BA 2024).

Tabelle 4-4: Übersicht der Ein- und Ausschlusskriterien für RCT mit dem zu bewertenden Arzneimittel gegenüber der zVT

Kategorie	Einschlusskriterien (E)		Ausschlusskriterien (A)	
Patientenpopulation	E1	Erwachsene Patienten mit PMR, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt	A1	Andere Patientenpopulation
Intervention	E2	Sarilumab (s.c. Injektion): 200 mg q2w	A2	Andere Dosierung oder Formulierung von Sarilumab oder andere Medikation
Vergleichstherapie	E3	Therapie nach ärztlicher Maßgabe unter Berücksichtigung von systemischen Glukokortikoiden und der Kombination von Glukokortikoiden mit Methotrexat	A3	Abweichende Vergleichsintervention
Endpunkte	E4	Mindestens ein patientenrelevanter Endpunkt aus den Kategorien <ul style="list-style-type: none"> • Mortalität • Morbidität • Gesundheitsbezogene Lebensqualität • Nebenwirkungen 	A4	Ungeeignete Endpunkte oder keine verwertbaren Ergebnisse zu patientenrelevanten Endpunkten
Studientyp	E5	RCT	A5	Andere Studien als RCT, tierexperimentelle Studien, in vitro-Studien
Studiendauer	E6	≥ 24 Wochen	A6	< 24 Wochen
Sprache	E7	Englisch oder Deutsch	A7	Andere als Englisch oder Deutsch
Publikationstyp	E8	Vollpublikation oder andere ausführliche Ergebnisberichte, z. B. Studienbericht, laufende Studie mit verfügbaren Zwischenergebnissen	A8	Mehrfachpublikationen ohne relevante Zusatzinformationen, Konferenz-Abstracts oder Poster
Studienstatus	E9	Studien, für die Ergebnisse vorliegen	A9	Noch nicht begonnene, laufende oder abgeschlossene Studien ohne publizierte Ergebnisse

Kategorie	Einschlusskriterien (E)	Ausschlusskriterien (A)
A: Ausschlusskriterium; E: Einschlusskriterium; PMR: Polymyalgia rheumatica; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); s.c.: Subkutan (Subcutaneous); zVT: Zweckmäßige Vergleichstherapie.		

4.2.3 Informationsbeschaffung

In den nachfolgenden Abschnitten ist zu beschreiben, nach welcher Methodik Studien identifiziert wurden, die für die Bewertung des medizinischen Nutzens und des medizinischen Zusatznutzens in dem in diesem Dokument bewerteten Anwendungsgebiet herangezogen werden. Dies bezieht sich sowohl auf publizierte als auch auf unpublizierte Studien. Die Methodik muss dazu geeignet sein, die relevanten Studien (gemäß den in Abschnitt 4.2.2 genannten Kriterien) systematisch zu identifizieren (systematische Literaturrecherche).

4.2.3.1 Studien des pharmazeutischen Unternehmers

Für die Identifikation der Studien des pharmazeutischen Unternehmers ist keine gesonderte Beschreibung der Methodik der Informationsbeschaffung erforderlich. Die vollständige Auflistung aller Studien, die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie aller Studien, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, erfolgt in den Abschnitten 4.3.1 und 4.3.2, jeweils im Unterabschnitt „Studien des pharmazeutischen Unternehmers“. Die Darstellung soll auf Studien mit Patienten in dem Anwendungsgebiet, für das das vorliegende Dokument erstellt wird, beschränkt werden.

4.2.3.2 Bibliografische Literaturrecherche

Die Durchführung einer bibliografischen Literaturrecherche ist erforderlich, um sicherzustellen, dass ein vollständiger Studienpool in die Bewertung einfließt.

Eine bibliografische Literaturrecherche muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine bibliografische Literaturrecherche immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die bibliografische Literaturrecherche soll mindestens in den Datenbanken MEDLINE (inklusive „in-process & other non-indexed citations“) und EMBASE sowie in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Datenbanken (z. B. CINAHL, PsycINFO etc.) durchgeführt werden.

Die Suche soll in jeder Datenbank einzeln und mit einer für die jeweilige Datenbank adaptierten Suchstrategie durchgeführt werden. Die Suchstrategien sollen jeweils in Blöcken, insbesondere getrennt nach Indikation, Intervention und ggf. Studientypen, aufgebaut werden. Wird eine Einschränkung der Strategien auf bestimmte Studientypen vorgenommen (z. B. randomisierte kontrollierte Studien), sollen aktuelle validierte Filter hierfür verwendet werden. Alle Suchstrategien sind in Anhang 4-A zu dokumentieren.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Datenbanken eine bibliografische Literaturrecherche durchgeführt wurde. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Sprach- oder Jahreseinschränkungen), und begründen Sie diese.

Die bibliografische Literaturrecherche nach RCT mit Sarilumab im vorliegenden AWG wurde am 22. November 2024 entsprechend den Vorgaben in den folgenden Datenbanken, jeweils über die Suchoberfläche Ovid, durchgeführt:

- Excerpta Medica dataBASE (EMBASE)
- Medical Literature Analysis and Retrieval System Online (MEDLINE) (inclusive “in-process & other non-indexed citations”)
- Cochrane Central Register of Controlled Trials

Für die Eingrenzung auf den Studientyp „Randomisierte, kontrollierte Studie“ wurde für EMBASE und MEDLINE der validierte Filter nach Wong verwendet (Wong 2006).

Die Suchstrategie wurde an die jeweilige Datenbank angepasst. Alle Suchvorgänge wurden nach den Kriterien „Population“, „Intervention“, „Comparison“, „Outcome“ und „Study design“ (PICOS) durchgeführt. Zeitliche Einschränkungen erfolgten nicht.

Detaillierte Angaben zu den Suchstrategien, dem Rechercheverlauf und den verwendeten Suchbegriffen finden sich in Anhang 4-A. Ergebnisse zu den Recherchen sind in Abschnitt 4.3.1.1.2 und Anhang 4-C dargestellt.

4.2.3.3 Suche in Studienregistern/ Studienergebnisdatenbanken

Eine Suche in öffentlich zugänglichen Studienregistern/ Studienergebnisdatenbanken ist grundsätzlich durchzuführen, um sicherzustellen, dass laufende Studien sowie abgeschlossene Studien auch von Dritten vollständig identifiziert werden und in Studienregistern / Studienergebnisdatenbanken vorliegende Informationen zu Studienmethodik und –ergebnissen in die Bewertung einfließen.

Eine Suche in Studienregistern/ Studienergebnisdatenbanken muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte

Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche in Studienregistern sowie Studienergebnisdatenbanken immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Die Suche soll mindestens in den Studienregistern/ Studienergebnisdatenbanken [clinicaltrials.gov](http://www.clinicaltrials.gov) (www.clinicaltrials.gov), EU Clinical Trials Register (EU-CTR, www.clinicaltrialsregister.eu), International Clinical Trials Registry Platform Search Portal (ICTRP Search Portal), Suchportal der WHO, Clinical Data Suchportal der European Medicines Agency (<https://clinicaldata.ema.europa.eu>) sowie dem Arzneimittel-Informationssystem (AMIS, <https://www.pharmnet-bund.de/dynamic/de/arsneimittel-informationssystem/index.html>) durchgeführt werden. Optional kann zusätzlich eine Suche in weiteren themenspezifischen Studienregistern / Studienergebnisdatenbanken (z. B. krankheitsspezifische Studienregister oder Studienregister einzelner pharmazeutischer Unternehmen) durchgeführt werden. Die Suche in Studienregistern/ Studienergebnisdatenbanken anderer pharmazeutischer Unternehmer ist insbesondere bei indirekten Vergleichen sinnvoll, wenn Studien zu anderen Arzneimitteln identifiziert werden müssen.

Die Suche soll in jedem Studienregister/ Studienergebnisdatenbank einzeln und mit einer für das jeweilige Studienregister/ Studienergebnisdatenbank adaptierten Suchstrategie durchgeführt werden. Die Suche soll abgeschlossene, abgebrochene und laufende Studien erfassen. Alle Suchstrategien sind in Anhang 4-B zu dokumentieren.

Für Clinical Data (Suchportal der European Medicines Agency) und das Arzneimittel-Informationssystem (AMIS) genügt hingegen die Suche nach Einträgen mit Ergebnisberichten zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Beschreiben Sie nachfolgend für alle durchgeführten Recherchen, in welchen Studienregistern/ Studienergebnisdatenbanken die Suche durchgeführt wurde. Begründen Sie dabei Abweichungen von den oben beschriebenen Vorgaben. Geben Sie auch an, wenn bei der Recherche generelle Einschränkungen vorgenommen wurden (z. B. Jahreseinschränkungen), und begründen Sie diese.

Die Suche nach RCT mit Sarilumab im vorliegenden AWG wurde am 22. November 2024 in den folgenden Studienregistern und Studienergebnisdatenbanken durchgeführt:

- [Clinicaltrials.gov](https://www.clinicaltrials.gov/) (<https://www.clinicaltrials.gov/>)
- EU Clinical Trials Register (EU-CTR) (<https://www.clinicaltrialsregister.eu>)

- International Clinical Trials Registry Platform (ICTRP) Search Portal der Weltgesundheitsorganisation (World Health Organization, WHO) (<https://trialsearch.who.int/>)
- Arzneimittelinformationssystem (AMIce), Datenbank des Bundesinstituts für Arzneimittel und Medizinprodukte (BfArM) (<https://portal.dimdi.de/>)

Für die Suche kam für jedes Studienregister bzw. jede Studienergebnisdatenbank eine angepasste Suchstrategie zum Einsatz. Die Recherche für AMIce beschränkte sich entsprechend den Vorgaben auf die Suche nach Einträgen mit Ergebnisberichten, die bereits anderweitig (z. B. über die bibliografische Literaturrecherche und Studienregisterrecherche) identifiziert wurden. Die Dokumentation der Suchstrategie in dieser Datenbank war nicht erforderlich. Weitere Einschränkungen der Suche wurden nicht vorgenommen.

Eine detaillierte Beschreibung der Studienregisterrecherche findet sich in Anhang 4-B. Die Ergebnisse der Suche in den Studienregistern sind in Abschnitt 4.3.1.1.3 und Anhang 4-D abgebildet.

4.2.3.4 Suche auf der Internetseite des G-BA

Die Internetseite des G-BA ist grundsätzlich zu durchsuchen, um sicherzustellen, dass alle vorliegenden Daten zu Studienmethodik und –ergebnissen von relevanten Studien in die Bewertung einfließen.

Auf der Internetseite des G-BA werden Dokumente zur frühen Nutzenbewertung nach §35a SGB V veröffentlicht. Diese enthalten teilweise anderweitig nicht veröffentlichte Daten zu Studienmethodik und –ergebnissen¹. Solche Daten sind dabei insbesondere in den Modulen 4 der Dossiers pharmazeutischer Unternehmer, in IQWiG-Nutzenbewertungen sowie dem Beschluss des G-BA einschließlich der Tragenden Gründe und der Zusammenfassenden Dokumentation zu erwarten.

Die Suche auf der Internetseite des G-BA muss für RCT mit dem zu bewertenden Arzneimittel (Abschnitt 4.3.1) immer durchgeführt werden. Für indirekte Vergleiche auf Basis von RCT (Abschnitt 4.3.2.1), nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) sowie weitere Untersuchungen (Abschnitt 4.3.2.3) muss eine Suche auf der G-BA Internetseite immer dann durchgeführt werden, wenn auf Basis solcher Studien der medizinische Zusatznutzen bewertet wird. Die Suche ist dann sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie durchzuführen. Es genügt die Suche nach Einträgen zu Studien, die bereits anderweitig (z.B. über die bibliografische Literaturrecherche und

¹ Köhler M, Haag S, Biester K, Brockhaus AC, McGauran N, Grouven U, Kölsch H, Seay U, Hörn H, Moritz G, Staack K, Wieseler B. Information on new drugs at market entry: retrospective analysis of health technology assessment reports, journal publications, and registry reports. *BMJ* 2015;350:h796

Studienregistersuche) identifiziert wurden. Eine Dokumentation der zugehörigen Suchstrategie ist nicht erforderlich.

Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Beschreiben Sie nachfolgend das Vorgehen für die Suche. Benennen Sie die Wirkstoffe und die auf der Internetseite des G-BA genannten zugehörigen Vorgangsnummern, zu denen Sie eine Suche durchgeführt haben.

Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Die Suche nach RCT mit Sarilumab im vorliegenden AWG wurde am 22. November 2024 entsprechend den Vorgaben auf der Website des G-BA (<https://www.g-ba.de/informationen/nutzenbewertung/>) durchgeführt. Dabei wurde nach Nutzenbewertungen entsprechend der (frühen) Nutzenbewertung nach § 35a Sozialgesetzbuch (SGB) V gesucht.

4.2.3.5 Selektion relevanter Studien

Beschreiben Sie das Vorgehen bei der Selektion relevanter Studien aus dem Ergebnis der in den Abschnitten 4.2.3.2, 4.2.3.3 und 4.2.3.4 beschriebenen Rechenschritte. Begründen Sie das Vorgehen, falls die Selektion nicht von zwei Personen unabhängig voneinander durchgeführt wurde.

Die Bewertung der in der bibliografischen Literaturrecherche, in Studienregister- / Studien- ergebnisdatenbanken und auf der Internetseite des G-BA identifizierten Publikationen erfolgte zunächst auf Titel- bzw. Abstract- und anschließend auf Volltextebene anhand der in Abschnitt 4.2.2 erläuterten Ein- und Ausschlusskriterien (Tabelle 4-4). Zwei Gutachter nahmen die Bewertung unabhängig voneinander vor, dabei wurden eventuelle Abweichungen in der Bewertung bis zu einer Konsensfindung diskutiert. Der Selektionsprozess wurde elektronisch dokumentiert. Die ausgeschlossenen Dokumente und Studien mit Ausschlussgrund befinden sich in Anhang 4-C und Anhang 4-D.

4.2.4 Bewertung der Aussagekraft der Nachweise

Zur Bewertung der Aussagekraft der im Dossier vorgelegten Nachweise sollen Verzerrungsaspekte der Ergebnisse für jede eingeschlossene Studie beschrieben werden, und zwar separat für jeden patientenrelevanten Endpunkt. Dazu sollen insbesondere folgende endpunktübergreifende (A) und endpunktspezifische (B) Aspekte systematisch extrahiert werden (zur weiteren Erläuterung der einzelnen Aspekte siehe Bewertungsbogen in Anhang 4-F):

A: Verzerrungsaspekte der Ergebnisse auf Studienebene

- Erzeugung der Randomisierungssequenz (*bei randomisierten Studien*)
- Verdeckung der Gruppenzuteilung (*bei randomisierten Studien*)
- zeitliche Parallelität der Gruppen (*bei nicht randomisierten vergleichenden Studien*)
- Vergleichbarkeit der Gruppen bzw. Berücksichtigung prognostisch relevanter Faktoren (*bei nicht randomisierten vergleichenden Studien*)
- Verblindung des Patienten sowie der behandelnden Personen
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

B: Verzerrungsaspekte der Ergebnisse auf Endpunktebene

- Verblindung der Endpunkterheber
- Umsetzung des ITT-Prinzips
- ergebnisgesteuerte Berichterstattung
- sonstige Aspekte

Für randomisierte Studien soll darüber hinaus das Verzerrungspotenzial bewertet und als „niedrig“ oder „hoch“ eingestuft werden. Ein niedriges Verzerrungspotenzial liegt dann vor, wenn mit großer Wahrscheinlichkeit ausgeschlossen werden kann, dass die Ergebnisse relevant verzerrt sind. Unter einer relevanten Verzerrung ist zu verstehen, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

Eine zusammenfassende Bewertung der Verzerrungsaspekte soll nicht für nicht randomisierte Studien erfolgen.

Für die Bewertung eines Endpunkts soll für randomisierte Studien zunächst das Verzerrungspotenzial endpunktübergreifend anhand der unter A aufgeführten Aspekte als „niedrig“ oder „hoch“ eingestuft werden. Falls diese Einstufung als „hoch“ erfolgt, soll das Verzerrungspotenzial für den Endpunkt in der Regel auch als „hoch“ bewertet werden, Abweichungen hiervon sind zu begründen. Ansonsten sollen die unter B genannten endpunktspezifischen Aspekte Berücksichtigung finden.

Eine Einstufung des Verzerrungspotenzials des Ergebnisses für einen Endpunkt als „hoch“ soll nicht zum Ausschluss der Daten führen. Die Klassifizierung soll vielmehr der Diskussion heterogener Studienergebnisse und der Einschätzung der Aussagekraft der Nachweise dienen. Für nicht randomisierte Studien können für solche Diskussionen einzelne Verzerrungsaspekte herangezogen werden.

Beschreiben Sie die für die Bewertung der Verzerrungsaspekte und des Verzerrungspotenzials eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Die Bewertung der Aussagekraft der Nachweise erfolgte auf Studienebene und separat für jeden patientenrelevanten Endpunkt auf Basis des klinischen Studienberichts zur Studie SAPHYR. Die Extraktion und Dokumentation wurden unter Verwendung von Anhang 4-F durchgeführt:

Das Verzerrungspotenzial auf Studienebene (A) wurde nach den Vorgaben der Modulvorlage für die endpunktübergreifenden bzw. endpunktspezifischen Aspekte bewertet und als „niedrig“ oder „hoch“ eingestuft.

Ausgehend von der Beurteilung des Verzerrungspotenzials auf Studien- und Endpunktebene erfolgte zunächst die Graduierung des Ausmaßes der Ergebnissicherheit auf Einzelstudien- und Endpunktebene und im Anschluss daran die Klassifizierung der Wahrscheinlichkeit des Zusatznutzens für jeden patientenrelevanten Endpunkt in „Beleg“, „Hinweis“, „Anhaltspunkt“ oder „kein nachgewiesener Zusatznutzen“ gemäß den Angaben im Methodenpapier des IQWiG (IQWiG 2023).

4.2.5 Informationssynthese und -analyse

4.2.5.1 Beschreibung des Designs und der Methodik der eingeschlossenen Studien

Das Design und die Methodik der eingeschlossenen Studien soll in den Abschnitten 4.3.1 und 4.3.2, jeweils in den Unterabschnitten „Charakteristika der in die Bewertung eingeschlossenen Studien“ und den dazugehörigen Anhängen, dargestellt werden. Die Darstellung der Studien soll für randomisierte kontrollierte Studien mindestens die Anforderungen des CONSORT-Statements erfüllen (Items 2b bis 14, Informationen aus dem CONSORT-Flow-Chart)². Die Darstellung nicht randomisierter Interventionsstudien und epidemiologischer Beobachtungsstudien soll mindestens den Anforderungen des TREND-³ bzw. STROBE-Statements⁴ folgen. Design und Methodik weiterer Untersuchungen sollen gemäß den verfügbaren Standards dargestellt werden.

Beschreiben Sie, nach welchen Standards und mit welchen Informationen (Items) Sie das Design und die Methodik der eingeschlossenen Studien in Modul 4 dargestellt haben. Begründen Sie Abweichungen von den oben beschriebenen Vorgaben.

Die Charakteristika der in die Bewertung eingeschlossenen, in Abschnitt 4.3.1.2 beschriebenen RCT SAPHYR wurden anhand der Items 2b bis 14b des CONSORT-Statements dargestellt (Anhang 4-E). Studienziele und -design, Patientenpopulation, Interventionen und Zielkriterien sind in den Items 2b bis 6, Fallzahlbestimmung, Randomisierung, Verblindung und statistische Methoden in den Items 7 bis 12 und Patientenfluss, inklusive Flow-Chart, ferner Informationen zu Aufnahme und Behandlungsdauer der Patienten sowie zum Studienende in den Items 13 bis 14b abgebildet.

² Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c332.

³ Des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Publ Health* 2004; 94(3): 361-366.

⁴ Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147(8): 573-577.

4.2.5.2 Gegenüberstellung der Ergebnisse der Einzelstudien

Die Ergebnisse der einzelnen Studien sollen in den Abschnitten 4.3.1 und 4.3.2 in den entsprechenden Unterabschnitten zunächst für jede eingeschlossene Studie separat dargestellt werden. Die Darstellung soll die Charakteristika der Studienpopulationen sowie die Ergebnisse zu allen in den eingeschlossenen Studien berichteten patientenrelevanten Endpunkten (Verbesserung des Gesundheitszustands, Verkürzung der Krankheitsdauer, Verlängerung des Überlebens, Verringerung von Nebenwirkungen, Verbesserung der Lebensqualität) umfassen. Anforderungen an die Darstellung werden in den Unterabschnitten beschrieben.

Benennen Sie die Patientencharakteristika und patientenrelevanten Endpunkte, die in den relevanten Studien erhoben wurden. Begründen Sie, wenn Sie von den oben benannten Vorgaben abgewichen sind. Beschreiben Sie für jeden Endpunkt, warum Sie ihn als patientenrelevant einstufen, und machen Sie Angaben zur Validität des Endpunkts (z. B. zur Validierung der eingesetzten Fragebögen). Geben Sie für den jeweiligen Endpunkt an, ob unterschiedliche Operationalisierungen innerhalb der Studien und zwischen den Studien verwendet wurden. Benennen Sie die für die Bewertung herangezogene(n) Operationalisierung(en) und begründen Sie die Auswahl. Beachten Sie bei der Berücksichtigung von Surrogatendpunkten Abschnitt 4.5.4.

Sofern zur Berechnung von Ergebnissen von Standardverfahren und –software abgewichen wird (insbesondere beim Einsatz spezieller Software oder individueller Programmierung), sind die Berechnungsschritte und ggf. verwendete Software explizit abzubilden. Insbesondere der Programmcode ist in lesbarer Form anzugeben.

Patientencharakteristika

Die Analysepopulation umfasst erwachsene Patienten mit PMR, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt. Tabelle 4-5 und Tabelle 4-6 geben einen Überblick über die in der Studie SAPHYR erfassten Patientencharakteristika zu Baseline.

Tabelle 4-5: Demografische Charakterisierung der Patientenpopulation zu Baseline

Kategorie	Ausprägung
Alter (Jahre)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Altersgruppe (Jahre), n (%)	<ul style="list-style-type: none"> • ≥ 50 - < 65 • ≥ 65 - < 75 • ≥ 75 - < 85 • ≥ 85
Geschlecht, n (%)	<ul style="list-style-type: none"> • Männlich • Weiblich
Abstammung, n (%)	<ul style="list-style-type: none"> • Weiß • Asiatisch
Bevölkerungsgruppenzugehörigkeit, n (%)	<ul style="list-style-type: none"> • Hispanisch oder lateinamerikanisch • Weder hispanisch noch lateinamerikanisch

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Kategorie	Ausprägung
Geografische Region, n (%)	<ul style="list-style-type: none"> • Westliche Länder • Südamerika • Rest der Welt
Körpergewicht (kg)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Gewichtsguppe (kg), n (%)	<ul style="list-style-type: none"> • < 60 • ≥ 60 - < 100 • ≥ 100
BMI (kg/m ²)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
BMI-Gruppe (kg/m ²), n (%)	<ul style="list-style-type: none"> • < 25 • ≥ 25 - < 30 • ≥ 30
Raucherstatus, n (%)	<ul style="list-style-type: none"> • Nie • Aktiv • Ehemals
Alkoholgenuss ^a , n (%)	<ul style="list-style-type: none"> • Nie • Gelegentlich • Mindestens monatlich • Mindestens wöchentlich • Mindestens täglich
<p>a: Innerhalb der letzten 12 Monate. BMI: Body Mass Index; Max: Maximum; Min: Minimum; MW: Mittelwert; n: Anzahl der Patienten mit Ausprägung; Q: Quartil; SD: Standardabweichung (Standard Deviation).</p>	

Tabelle 4-6: Krankheitsspezifische Charakterisierung der Patientenpopulation zu Baseline

Kategorie	Ausprägung
Morgensteifigkeit, n (%)	<ul style="list-style-type: none"> • Nein • Ja <ul style="list-style-type: none"> ○ ≤ 45 min ○ > 45 min
Morgensteifigkeit (min)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Schulerschmerzen, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Eingeschränkte Schulterbeweglichkeit, n (%)	<ul style="list-style-type: none"> • Nein • Ja <ul style="list-style-type: none"> ○ Arm-Anheben über Schulterhöhe ○ Arm-Anheben unter Schulterhöhe ○ Arm-Anheben bis Schulter ○ Kein Arm-Anheben

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Kategorie	Ausprägung
Hüftschmerzen, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Eingeschränkte Hüftbeweglichkeit, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Konstitutionelle Symptome, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Beteiligung anderer Gelenke ^a , n (%)	<ul style="list-style-type: none"> • Ja • Nein
Andere Merkmale, die auf einen PMR-Schub hindeuten, n (%)	<ul style="list-style-type: none"> • Ja • Nein
PMR-Diagnose mittels Ultraschall, n (%)	<ul style="list-style-type: none"> • Ja • Nein
Prednison-Dosis (oder Äquivalent) während PMR-Schub (mg)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Patienten mit vorherigen PMR-Schüben, n (%) ^b	<ul style="list-style-type: none"> • ≤ 6 Monate vor Screening • 6 – 12 Monate vor Screening • > 12 Monate vor Screening
Anzahl vorheriger PMR-Schübe pro Patient	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Prednison-Dosis (oder Äquivalent) für PMR zu Baseline (mg/Tag) ^c	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Höchste CS-Dosis ^d für PMR 24 Wochen vor Screening bis Baseline (mg/Tag)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Dauer des CS-Gebrauchs ^e für PMR 24 Wochen vor Screening bis Baseline (Tage)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Dauer der PMR von Diagnose bis Baseline (Tage)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
CRP zu Baseline (mg/l)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
ESR zu Baseline (mm/h)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
Arztbeurteilung der Krankheitsaktivität (MD-VAS)	<ul style="list-style-type: none"> • n • MW (SD) • Median (Min : Max) • Q1 : Q3
GTI zu Baseline (0-538)	<ul style="list-style-type: none"> • n

Kategorie	Ausprägung
	<ul style="list-style-type: none"> • MW (SD) • Median (Min : Max) • Q1 : Q3
Vorherige biologische DMARD, n (%)	<ul style="list-style-type: none"> • Adalimumab • Tocilizumab
Vorherige nicht-biologische DMARD, n (%)	<ul style="list-style-type: none"> • Methotrexat • Methotrexat-Natrium • Leflunomid • Azathioprin • Hydroxychloroquin • Hydroxychloroquin-Sulfat
<p>a: Andere entzündete Gelenke (z. B. Hände, Handgelenke, Knie etc.) mit Bezug zur Krankheitsaktivität der PMR.</p> <p>b: Mehrere vorherige Schübe pro Patient und damit Zählung in mehr als einer Kategorie möglich.</p> <p>c: Basierend auf der letzten Dosis vor der Baseline-Visite.</p> <p>d: Einschließlich Prednison oder Äquivalent.</p> <p>e: Gesamtdauer des CS-Gebrauchs ohne Dosierungsunterbrechungen.</p> <p>CRP: C-reaktives Protein; CS: Corticosteroids; DMARD: Krankheitsmodifizierende Antirheumatika (Disease-Modifying Anti-Rheumatic Drugs); ESR: Erythrozyten-Sedimentationsrate; GTI: Glukokortikoid-Toxizitätsindex; Max: Maximum; MD-VAS: Visuelle Analogskala (Arztbeurteilung); Min: Minimum; MW: Mittelwert; n: Anzahl der Patienten mit Ausprägung; PMR: Polymyalgia rheumatica; Q: Quartil; SD: Standardabweichung (Standard Deviation).</p>	

Patientenrelevante Endpunkte

Entsprechend § 2 Abs. 3 der AM-NutzenV besteht der Nutzen eines Arzneimittels in einem patientenrelevanten therapeutischen Effekt (BMJ 2023). Als patientenrelevant erachten G-BA und IQWiG übereinstimmend die Kategorien Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen (G-BA 2025; IQWiG 2023). Dabei gilt es zu beurteilen, wie ein Patient sich fühlt, seine Funktionen und Aktivitäten wahrnehmen kann und ob bzw. wie lange er überlebt.

Die zur Nutzenbewertung im Rahmen des vorliegenden Dossiers herangezogenen patientenrelevanten Endpunkte sind in Tabelle 4-7 abgebildet. Eine ausführliche Beschreibung der Patientenrelevanz, Operationalisierung und Validität findet sich im anschließenden Fließtext.

Tabelle 4-7: Patientenrelevante Endpunkte zum Nachweis des Zusatznutzens von Sarilumab im vorliegenden AWG

Studienendpunkt	Mortalität	Morbidität	Gesundheits- bezogene Lebensqualität	Nebenwirkungen
Anhaltende Remission Woche 52, inkl. 4 Einzelkomponenten ^a		■		
Kumulative CS-Gesamtdosis		■		
Zeit bis zum ersten PMR-Schub nach klinischer Remission		■		
PMR-AS, inkl. fünf Einzelkomponenten ^b		■		
FACIT-F		■		
EQ-5D VAS		■		
HAQ-DI		■		
SF-36v2			■	
Sicherheit <ul style="list-style-type: none"> • UE gesamt • SUE • Schwere UE • Therapieabbruch aufgrund von UE • UE von besonderem Interesse 	■ ^c			■

a: Einzelkomponenten: Remission spätestens zu Woche 12, kein Krankheitsschub von Woche 12 bis Woche 52, anhaltende CRP-Reduktion (auf < 10 mg/l, ohne nachfolgende Anstiege auf ≥ 10 mg/l) von Woche 12 bis Woche 52, erfolgreiches Prednison-Ausschleichen von Woche 12 bis Woche 52.
b: Einzelkomponenten: CRP, VAS (Patienten- und Prüfarztbeurteilung), Dauer der Morgensteifigkeit, Beweglichkeit der oberen Gliedmaßen.
c: Beurteilung anhand von UE mit Todesfolge.

AWG: Anwendungsgebiet; CRP: C-reaktives Protein; CS: Corticosteroids; EQ-5D VAS: European Quality of Life 5 Dimensions Visuelle Analogskala; FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI: Health Assessment Questionnaire-Disability Index; PMR: Polymyalgia rheumatica; PMR-AS: PMR-Aktivitätsscore; SF-36v2: 36-Item Short-Form Health Survey Version 2; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala.

Mortalität

Gesamtmortalität

Patientenrelevanz

Die Verlängerung des Überlebens stellt eine wesentliche Zielsetzung einer Arzneimitteltherapie dar (BMJ 2023; G-BA 2025; IQWiG 2023). Da die PMR keine tödlich verlaufende Erkrankung ist, sind Studien im Indikationsgebiet nicht darauf ausgelegt, Mortalität zu untersuchen. Dennoch ist das Überleben der Patienten (bzw. eine mögliche Verlängerung des Überlebens) gemäß der Verfahrensordnung (VerfO) des G-BA grundsätzlich ein patientenrelevanter Endpunkt.

Operationalisierung

Für die vorliegende Nutzenbewertung werden Todesfälle im Rahmen der standardisierten Erhebung unerwünschter Ereignisse erfasst.

Validität

Endpunkte aus der Kategorie Mortalität unterliegen keiner subjektiven Bewertung und sind demzufolge grundsätzlich als valide anzusehen.

Morbidität

Anhaltende Remission zu Woche 52

Patientenrelevanz

Die Verbesserung des Gesundheitszustands stellt gemäß AM-NutzenV und Verfo des G-BA einen patientenrelevanten therapeutischen Arzneimitteleffekt dar (BMJ 2023; G-BA 2025). Eine anhaltende Remission der PMR spiegelt nicht nur eine Verbesserung des Gesundheitszustands wider, sondern ist darüber hinaus mit einer Steigerung der Lebensqualität assoziiert. Die Lebensqualität wird in der Nutzenbewertung medizinischer Interventionen ebenfalls als unmittelbar patientenrelevante Zielgröße beschrieben (BMJ 2023; G-BA 2025; IQWiG 2023). Die Patientenrelevanz des Endpunkts ist somit gegeben.

Operationalisierung

Die anhaltende Remission zu Woche 52 setzt sich zusammen aus den 4 Einzelkomponenten:

1. Remission spätestens zu Woche 12,
2. Kein Krankheitsschub von Woche 12 bis Woche 52,
3. Anhaltende C-reaktives Protein (CRP)-Reduktion von Woche 12 bis Woche 52,
4. Erfolgreiches Prednison-Ausschleichen von Woche 12 bis Woche 52.

Ad 1. Remission wurde definiert als Auflösung der Zeichen und Symptome der PMR und CRP-Normalisierung.

Die Beurteilung der Zeichen und Symptome der PMR beinhaltet unter anderem folgende Kriterien:

- Morgensteifigkeit und/oder Schmerzen in Nacken, Schulter und/oder Beckengürtel,
- Eingeschränkte Bewegungsfähigkeit in Schultern und/oder Beckengürtel,
- Konstitutionelle Symptome wie Fatigue, Gewichtsverlust und leichtes Fieber,
- Andere Anzeichen, die nach Ansicht des Prüfarztes auf einen PMR-Schub hinweisen.

Ad 2. Krankheitsschub wurde definiert als Wiederauftreten von Zeichen und Symptomen oder als Anstieg der ESR, im Zusammenhang mit aktiver PMR, die eine Erhöhung der CS-Dosis erforderlich machten. Erhöhung der CS-Dosis wurde definiert als:

- Jede Dosiserhöhung während des protokollgerechten Prednison-Ausschleichens,
- Erneute CS-Gabe nach Beendigung des protokollgerechten Ausschleichens.

Ad 3. Anhaltende CRP-Reduktion wurde definiert als Normalisierung des CRP auf < 10 mg/l ohne sukzessive Anstiege auf ≥ 10 mg/l.

Ad 4. Erfolgreiches Prednison-Ausschleichen wurde definiert als ausbleibender Bedarf an Zusatzmedikation. Der Gebrauch von zusätzlichem Prednison mit einer kumulativen Dosis von ≤ 100 mg (oder Äquivalent), d. h., über die in der Ausschleichphase vorgesehene Dosis hinaus, beispielsweise zur Behandlung von nicht im Zusammenhang mit der PMR stehenden UE, war möglich.

Die Ergebnisse werden in Form von Responderanalysen dargestellt. Odds Ratio (OR), Relatives Risiko (RR) und Risikodifferenz (RD) mit den jeweils zugehörigen 95 %-Konfidenzintervallen (KI) und p-Werten wurden für die anhaltende Remission zu Woche 52 auf Basis aller 4 Einzelkomponenten und für die Einzelkomponenten mittels logistischer Regression berechnet.

Die Analyse basierte auf den Daten der Intention-To-Treat (ITT)-Population, die alle randomisierten Patienten einschloss. Für die Primäranalyse der anhaltenden Remission zu Woche 52 wurden Patienten, die die Studie vor Woche 52 beendeten und vor dem Abbruch keinen Schub erlitten, als Non-Responder imputiert. Fehlende Werte der Einzelkomponenten wurden nicht imputiert.

Validität

Eine vollständige und anhaltende Remission ist das primäre Behandlungsziel von Patienten mit PMR. Die Validität des Endpunkts ist durch die Beurteilung der Einzelkomponenten nach objektiven Kriterien gegeben.

Kumulative CS-Gesamtdosis über 52 Wochen

Patientenrelevanz

Corticosteroide sind die Grundlage in der Behandlung der PMR (Buttgereit 2018; Caporali 2004; Dejaco 2011). Die mit der Einnahme von Corticosteroiden assoziierte Toxizität (BPAC 2023; Liew 2018) bedeutet jedoch für die Patienten eine erhebliche Beeinträchtigung des körperlichen Wohlbefindens und der gesundheitsbezogenen Lebensqualität (Bridgewater 2023; Sullivan 2017). Häufigkeit und Schwere der Nebenwirkungen korrelieren sowohl mit der Dauer der Therapie als auch mit der Höhe der Dosis. Eine Reduktion der CS-Dosis bzw. eine möglichst niedrige Gesamtdosis ist mit der daraus resultierenden geringeren Toxizität von unmittelbarer Patientenrelevanz.

Operationalisierung

Die kumulative CS-Gesamtdosis wurde als Differenz zwischen tatsächlicher und erwarteter Gesamtdosis nach 52 Behandlungswochen zusammengefasst. Der erwarteten kumulativen Dosis liegt das protokollgerechte Prednison-Ausschleichen im Behandlungsverlauf zugrunde, während die tatsächliche kumulative Dosis neben dem protokollgerechten Prednison-Ausschleichen auch die Gabe von add-on-Prednison bis Woche 12 und Notfallmedikation von Woche 12 bis 52 berücksichtigt.

Es wurden der Median, einschließlich Minimum und Maximum, mit 95 %-KI und dem zugehörigem p-Wert anhand des nichtparametrischen Wilcoxon-Rangsummentests bestimmt. Die Analyse basierte auf den Daten der ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten (Sanofi 2021a).

Validität

Die kumulative CS-Gesamtdosis zählt zu den Endpunkten, die im Rahmen klinischer Studien an Patienten mit PMR berichtet werden (Caporali 2004; Catanoso 2007; Gabriel 1997; Salvarani 2007; Weyand 1999). Sie markiert somit ein valides Kriterium insbesondere zur Beurteilung CS-bezogener Nebenwirkungen (Dejaco 2015; Gabriel 1997).

Zeit bis zum ersten PMR-Schub während der Behandlungsphase

Patientenrelevanz

Die Verbesserung des Gesundheitszustands stellt gemäß AM-NutzenV und Verfahrensordnung des G-BA einen patientenrelevanten therapeutischen Arzneimitteleffekt dar (BMJ 2023; G-BA 2025). Ein längeres Andauern der Remission und in der Konsequenz das Hinauszögern eines PMR-Schubs spiegelt nicht nur eine Verbesserung des Gesundheitszustands wider, sondern ist darüber hinaus mit einer längeren Zeitspanne bei besserer gesundheitsbezogener Lebensqualität sowie einer in der Summe geringeren CS-Dosis verbunden. Die Zeit bis zum Auftreten eines Krankheitsschubs wird auch von G-BA und IQWiG als patientenrelevant eingestuft (G-BA 2025; IQWiG 2023).

Operationalisierung

Die Zeit bis zum ersten PMR-Schub wurde definiert als Zeitspanne zwischen Randomisierung und Auftreten des ersten PMR-Schubs nach Erreichen einer klinischen Remission. Patienten, die keine Remission erreichten, wurden als Patienten mit Schub zur Randomisierung betrachtet. Patienten mit Remission, die keinen Schub entwickelten, wurden zum Behandlungsende bis Woche 52 zensiert.

Der Behandlungseffekt (Hazard Ratio (HR) mit zugehörigem zweiseitigem 95 %-KI und p-Wert) wurde durch eine Cox-Regression geschätzt. Neben Anzahl der Patienten mit Ausprägung wurde die mediane Dauer bis zum ersten Schub nach klinischer Remission, einschließlich 95 %-KI nach der Kaplan-Meier-Methode unter Berücksichtigung der Brookmeyer-Crowley-Methode (log-log-Transformation) ermittelt und zusätzlich grafisch dargestellt. Die Analyse basierte auf den Daten der ITT-Population, die alle randomisierten Patienten einschloss (Sanofi 2021a).

Validität

Die Zeit von der Remission bis zum PMR-Schub zählt zu den häufig analysierten Endpunkten im Rahmen klinischer Studien an Patienten mit PMR (Catanoso 2007; Dejaco 2011; Salvarani 2007). Die Validität des Endpunkts ist durch die Beurteilung anhand objektiver Kriterien wie der Erythrozyten-Sedimentationsrate (ESR) und dem PMR-AS gegeben.

PMR-Aktivitätsscore (PMR-AS)

Patientenrelevanz

Der PMR-AS bildet wesentliche Kriterien der PMR im Hinblick auf die Beurteilung des therapeutischen Effekts und der Krankheitsschwere ab und ist somit unmittelbar patientenrelevant.

Operationalisierung

Der PMR-AS dient der Messung der Krankheitsaktivität. Er wurde aus den EULAR-Response-Kriterien für PMR entwickelt und umfasst folgende Parameter:

- CRP (mg/dl),
- VAS - Schmerz, Patientenbeurteilung (skaliert auf den Bereich von 0 bis 10),
- VAS - Krankheitsaktivität, Prüfarztbeurteilung (skaliert auf den Bereich von 0 bis 10),
- Dauer der Morgensteifigkeit (min x 0,1),
- Beweglichkeit der oberen Gliedmaßen (EUL: 3 = keine, 2 = bis unterhalb Schultergürtel, 1 = bis Schultergürtel, 0 = bis oberhalb Schultergürtel).

Die Ergebnisse der einzelnen Parameter werden aufsummiert, wobei ein PMR-AS unter 7 eine niedrige, zwischen 7 und 17 eine mittlere und über 17 eine hohe Krankheitsaktivität anzeigt.

Die Erhebung des PMR-AS erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Zur Bewertung der Schmerzsymptomatik durch den Patienten anhand einer visuellen Analogskala (VAS) wurden die Ergebnisse der Fragebogenerhebung des HAQ-DI herangezogen, da sich beide VAS lediglich in ihrer Skalierung unterscheiden. Die Ergebnisse der Einzelkomponenten VAS – Schmerz, Patientenbeurteilung, und VAS – Krankheitsaktivität, Prüfarztbeurteilung, werden als Veränderung um $\geq -1,5$ Punkte gegenüber Baseline, entsprechend dem Responsekriterium $\geq 15\%$ der Skalenspannweite der VAS, dargestellt. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Die Ergebnisse für den globalen PMR-AS, die übrigen 3 Einzelkomponenten sowie zusätzlich für die beiden VAS werden als Veränderung zu Woche 52 gegenüber Baseline anhand eines Mixed Effects Model for Repeated Measures (MMRM) analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.

Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung des PMR-AS und der Einzelkomponenten sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.

Validität

Die Validität des Endpunkts PMR-AS konnte in Studien wiederholt gezeigt werden (Binard 2009; Leeb 2004; Leeb 2007; Steindl 2009).

FACIT-F

Patientenrelevanz

Fatigue zählt zu den häufigsten und am stärksten belastenden Symptomen bei schweren Erkrankungen (Butt 2013; Montan 2018) und ist somit in der Beurteilung des allgemeinen Gesundheitszustands durch den Patienten von unmittelbarer Patientenrelevanz. Die valide Erfassung der Veränderung des Gesundheitszustands gibt zuverlässig wieder, wie der Patient sich fühlt und inwiefern er am Alltags- und Sozialleben teilnehmen kann. Gemäß AM-NutzenV und Verfahrensordnung des G-BA ist die Verbesserung des Gesundheitszustands als patientenrelevant anzusehen (BMJ 2023; G-BA 2025). In der Niederschrift zum Beratungsgespräch betrachtet der G-BA das mittels FACIT-F erfasste Symptom Fatigue für die PMR als patientenrelevant (G-BA 2024).

Operationalisierung

Der Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) wurde als befindlichkeitsspezifische Subskala des generischen Fragebogens FACIT-General (FACIT-G) zur Erfassung des Erschöpfungsgrads der Patienten angewandt. Er beinhaltet 13 Fragen, mit denen der Grad der Erschöpfung in den vergangenen 7 Tagen über eine fünfstufige Likert-Skala von 0 bis 4 (0 = überhaupt nicht, 1 = ein wenig, 2 = mäßig, 3 = ziemlich, 4 = sehr) gemessen wird. Die individuellen Antwortwerte werden auf einen Gesamtwert von 0 bis 52 summiert, wobei die Antwortwerte der 11 negativ formulierten Fragen umgekehrt werden. Ein höherer Gesamtwert bedeutet einen besseren Gesundheitszustand.

Die Erhebung des Fragebogens erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Die Ergebnisse werden zum einen als Veränderung um $\geq 7,8$ Punkte, entsprechend dem Responsekriterium ≥ 15 % der Skalenspannweite des FACIT-F, berichtet. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Zum anderen werden die Ergebnisse als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.

Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung des FACIT-F sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.

Validität

Der FACIT-F ist ein international eingesetzter, in vielen Sprachen validierter patientenberichteter Endpunkt (Patient-Reported Outcome, PRO) zur Erfassung des allgemeinen Gesundheitszustands durch den Patienten (Cella 2005). Eine Erhebung an einer repräsentativen Stichprobe der deutschen Allgemeinbevölkerung bestätigte Reliabilität und Validität der deutschen Version der FACIT-F-Subskala (Montan 2018).

EQ-5D VAS

Patientenrelevanz

Die Beurteilung des allgemeinen Gesundheitszustands durch den Patienten ist von unmittelbarer Patientenrelevanz. Die valide Erfassung der Veränderung des Gesundheitszustands gibt zuverlässig wieder, wie der Patient sich fühlt und inwiefern er am Alltags- und Sozialleben teilnehmen kann. Gemäß AM-NutzenV und Verfo des G-BA ist die Verbesserung des Gesundheitszustands als patientenrelevant anzusehen (BMJ 2023; G-BA 2025).

Operationalisierung

Der European Quality of Life 5 Dimensions (EQ-5D) ist ein generischer Fragebogen zur Erfassung des Gesundheitszustands der Patienten.

Anhand der VAS des EQ-5D ordnet der Patient seinem aktuellen Gesundheitszustand einen Wert zwischen 0 (denkbar schlechtester Gesundheitszustand) und 100 (denkbar bester Gesundheitszustand) zu.

Die Erhebung erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Die Ergebnisse werden zum einen als Veränderung um ≥ 15 Punkte, entsprechend dem Responsekriterium $\geq 15\%$ der Skalenspannweite der EQ-5D VAS, berichtet. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Zum anderen werden die Ergebnisse als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.

Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung der EQ-5D VAS sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.

Validität

Der EQ-5D ist ein weit verbreitetes, in vielen Sprachen, darunter auch in Deutsch, validiertes PRO-Instrument zur Erfassung des allgemeinen Gesundheitszustands durch den Patienten (Herdman 2011; Janssen 2013). Auch der G-BA berücksichtigt den EQ-5D-Fragebogen regelhaft in seinen Nutzenbewertungsverfahren zur Einstufung des Zusatznutzens, beispielsweise bei Sarilumab (Anwendungsgebiet Rheumatoide Arthritis) (G-BA 2018). Insbesondere die VAS des EQ-5D wird vom G-BA als valides Beurteilungskriterium des Gesundheitszustands der Patienten gewertet (G-BA 2018, 2021).

HAQ-DI

Patientenrelevanz

Die Erfassung funktioneller Einschränkungen durch den Patienten ist bei der Selbstbeurteilung des Gesundheitszustands unmittelbar patientenrelevant. Die valide Erfassung der Veränderung der Körperfunktionen gibt zuverlässig wieder, wie der Patient sich fühlt und inwiefern er am Alltags- und Sozialleben teilnehmen kann. Gemäß AM-NutzenV und Verfahrensordnung des G-BA ist die Verbesserung des Gesundheitszustands und somit der funktionellen Kapazität als patientenrelevant anzusehen (BMJ 2023; G-BA 2025). Der mittels HAQ-DI vom Patienten berichtete körperliche Funktionsstatus wird vom G-BA in der Niederschrift zum Beratungsgespräch als patientenrelevant erachtet (G-BA 2024).

Operationalisierung

Der HAQ-DI wurde als spezifischer Fragebogen für rheumatische Erkrankungen zur Erfassung der körperlichen Mobilität bzw. zur Einschätzung der funktionellen Einschränkungen der Patienten anhand der folgenden 8 Kategorien angewandt:

- Ankleiden und Körperpflege,
- Aufstehen,
- Essen,
- Gehen,
- Hygiene,
- Erreichen von Gegenständen,
- Greifen,
- Andere Tätigkeiten.

Jede Kategorie beinhaltet Fragen zu 2 oder 3 spezifischen Aktivitäten, mit denen der Patient seine funktionellen Einschränkungen in den vergangenen 7 Tagen beurteilt. Darüber hinaus wird die Benutzung von Hilfsmitteln (7 Fragen) bzw. die Notwendigkeit fremder Hilfe (4 Fragen) angegeben. Die Beantwortung der Fragen erfolgt über eine vierstufige Likert-Skala von 0 bis 3 (0 = ohne jede Schwierigkeit, 1 = mit einigen Schwierigkeiten, 2 = mit großen Schwierigkeiten, 3 = nicht dazu in der Lage). Zur Berechnung des HAQ-DI werden die höchsten Punktwerte aus den 8 Kategorien gemittelt. Je höher der Index ausfällt, desto schwerer ist die Einschränkung im Alltag.

Neben dem Standard-Score beinhaltet der HAQ-DI die Beurteilung des Schmerzempfindens und der globalen Krankheitsaktivität durch den Patienten mittels VAS auf einer Skala von 0 (keine Schmerzen bzw. sehr gute Gesundheit) bis 100 (sehr starke Schmerzen bzw. sehr schlechte Gesundheit).

Die Erhebung des Fragebogens erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Die Ergebnisse für den Standard-Score und die Beurteilung der globalen Krankheitsaktivität werden zum einen als Veränderung um $\geq -0,45$ Punkte (Standard-Score) bzw. ≥ -15 Punkte (globale Krankheitsaktivität), entsprechend dem Responsekriterium $\geq 15\%$ der Skalenspannweite des HAQ-DI, berichtet. OR, RR und RD mit den zugehörigen 95 %-KI und

p-Werten wurden mittels logistischer Regression berechnet. Zum anderen werden die Ergebnisse sowohl für den Standard-Score als auch für die globale Krankheitsaktivität und die 8 Einzelkategorien als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.

Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung des HAQ-DI sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.

Die Ergebnisse des Schmerzempfindens durch den Patienten mittels VAS werden als eine der Einzelkomponenten des Endpunkts PMR-AS dargestellt. Dabei wurde die VAS auf den Bereich von 0 (keine Schmerzen) bis 10 (sehr starke Schmerzen) skaliert.

Validität

Der HAQ-DI wurde zur Erfassung der körperlichen Funktionseinschränkungen bei Patienten mit rheumatoider Arthritis entwickelt (Fries 1980) und in der Folgezeit in vielen weiteren Indikationen einschließlich PMR angewandt (Bruce 2003; Kalke 2000; Leung 2023). Er stellt ein weltweit häufig eingesetztes Selbstbeurteilungsinstrument dar, wurde von der englischen Originalversion in zahlreiche Sprachen übersetzt, darunter auch in Deutsch, und ist validiert.

Gesundheitsbezogene Lebensqualität

SF-36v2

Patientenrelevanz

Die Beurteilung der gesundheitsbezogenen Lebensqualität durch den Patienten ist von unmittelbarer Patientenrelevanz. Die valide Erfassung der Veränderung der gesundheitsbezogenen Lebensqualität gibt zuverlässig wieder, wie der Patient sich fühlt und inwiefern er am Alltags- und Sozialleben teilnehmen kann. Gemäß AM-NutzenV und Verfahrensordnung des G-BA ist die Verbesserung der gesundheitsbezogenen Lebensqualität als patientenrelevant anzusehen (BMJ 2023; G-BA 2025). Der G-BA bestätigte dies in der finalen Niederschrift zum Beratungsgespräch (G-BA 2024).

Operationalisierung

Der SF-36 wurde als generischer Fragebogen zur Erfassung der gesundheitsbezogenen Lebensqualität der Patienten anhand der beiden Hauptdimensionen PCS und MCS angewandt.

Die PCS wird aus den folgenden Subskalen gebildet:

- Körperliche Funktionsfähigkeit,
- Körperliche Rollenfunktion,
- Körperliche Schmerzen,

- Allgemeine Gesundheitswahrnehmung.

Die MCS wird aus den folgenden Subskalen gebildet:

- Vitalität,
- Soziale Funktionsfähigkeit,
- Emotionale Rollenfunktion,
- Psychisches Wohlbefinden.

Pro Subskala ergeben 2 bis 10 Fragen insgesamt 36 Fragen, von denen eine Frage zur Veränderung der allgemeinen Gesundheitswahrnehmung in den vergangenen 12 Monaten in der Auswertung unberücksichtigt bleibt. Alle Fragen werden mit Hilfe einer drei- bis sechsstufigen Likert-Skala beantwortet und die erhaltenen Werte für jede Subskala in einen Bereich zwischen 0 (schlechteste Lebensqualität) und 100 (beste Lebensqualität) transformiert. PCS und MCS werden auf einen Mittelwert von 50 und eine Standardabweichung von 10 transformiert. Höhere Werte bedeuten eine bessere gesundheitsbezogene Lebensqualität.

Die Erhebung des Fragebogens erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Die Ergebnisse werden zum einen als Veränderung um ≥ 10 Punkte berichtet. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Zum anderen werden die Ergebnisse als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.

Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung der beiden Summenskalen des SF-36v2 sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.

Validität

Der SF-36 ist ein weltweit eingesetztes generisches PRO-Instrument zur Messung der gesundheitsbezogenen Lebensqualität (Ware, JR 1992; Ware, JR 2000). Er wurde in mehr als 100 Sprachen übersetzt (Methode der Rückübersetzung) und validiert, darunter auch in Deutsch (Bullinger 1995). Die Version 2 des SF-36 ist ebenfalls in über 100 Übersetzungen verfügbar; die deutsche Version 2 wurde erstmals im Jahr 2002 eingesetzt (Ellert 2013; Morfeld 2008).

Nebenwirkungen (unerwünschte Ereignisse, UE)

Patientenrelevanz

UE als Maß für die Verträglichkeit einer Therapie (therapiebedingte Morbidität) werden generell als patientenrelevanter Endpunkt betrachtet, da sie sich unmittelbar auf den Gesundheitszustand eines Patienten auswirken (BMJ 2023; G-BA 2025; IQWiG 2023).

Operationalisierung

Es wurden alle in der Behandlungsphase aufgetretenen UE vom Zeitpunkt der ersten bis 60 Tage nach der letzten Verabreichung der Studienmedikation erfasst. Die Kodierung der UE erfolgte anhand der MedDRA-Version 24.0 hierarchisch von der entsprechenden Systemorganklasse (System Organ Class, SOC) zum bevorzugten Begriff (Preferred Term, PT). Schwerwiegende UE (SUE) oder UE von besonderem Interesse, die zum Studienende weiter bestanden, wurden bis zur Auflösung oder Stabilisierung weiter beobachtet und dokumentiert.

Folgende sicherheitsrelevante Endpunkte wurden berücksichtigt:

- UE gesamt,
- Schwere UE,
- SUE,
- UE, die zum Therapieabbruch führten,
- UE von besonderem Interesse,
 - Schwangerschaft einer Studienpatientin oder Partnerin eines Studienpatienten,
 - Symptomatische Überdosierung der Studienmedikation,
 - ALT-Anstieg auf $\geq 3 \times$ ULN,
 - Klinisch relevante Infektionen,
 - Abnormale ALT- oder ANC-Werte oder Thrombozytopenie, die zum Therapieabbruch führten.

Die Ergebnisse der sicherheitsrelevanten Endpunkte werden in Form von Responderanalysen dargestellt. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Darüber hinaus werden UE gesamt, die bei mindestens 10 % der Patienten in einem Studienarm auftraten, sowie alle schweren UE und SUE, die bei mindestens 5 % der Patienten in einem Studienarm auftraten, in Form von Responderanalysen nach SOC und PT berichtet. UE, die zum Therapieabbruch führten, und UE von besonderem Interesse werden zusätzlich deskriptiv nach SOC und PT berichtet.

Die Analyse basierte auf den Daten der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten.

Validität

Zur Bewertung der Patientensicherheit, einschließlich klinischer und laborbezogener UE, während der Studie stand ein unabhängiger Sicherheitsbeauftragter zur Verfügung (Sanofi 2021b). Da eine subjektive Beurteilung der UE durch die Prüfarzte somit ausgeschlossen war und ihre Dokumentation darüber hinaus internationalen Standards wie MedDRA folgte, ist die Validität des Endpunkts Nebenwirkungen sichergestellt.

4.2.5.3 Meta-Analysen

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse quantitativ zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z.B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Für Meta-Analysen soll die im Folgenden beschriebene Methodik eingesetzt werden.

Für die statistische Auswertung sollen primär die Ergebnisse aus Intention-to-treat-Analysen, so wie sie in den vorliegenden Dokumenten beschrieben sind, verwendet werden. Die Meta-Analysen sollen in der Regel auf Basis von Modellen mit zufälligen Effekten nach der Knapp-Hartung-Methode mit der Paule-Mandel-Methode zur Heterogenitätsschätzung⁵ erfolgen. Im Fall von sehr wenigen Studien ist die Heterogenität nicht verlässlich schätzbar. Liegen daher weniger als 5 Studien vor, ist auch die Anwendung eines Modells mit festem Effekt oder eine qualitative Zusammenfassung in Betracht zu ziehen. Kontextabhängig können auch alternative Verfahren wie z. B. Bayes'sche Verfahren oder Methoden aus dem Bereich der generalisierten linearen Modelle in Erwägung gezogen werden. Falls die für eine Meta-Analyse notwendigen Schätzer für Lage und Streuung in den Studienunterlagen nicht vorliegen, sollen diese nach Möglichkeit aus den vorhandenen Informationen eigenständig berechnet beziehungsweise näherungsweise bestimmt werden.

Für kontinuierliche Variablen soll die Mittelwertdifferenz, gegebenenfalls standardisiert mittels Hedges' g, als Effektmaß eingesetzt werden. Bei binären Variablen sollen Meta-Analysen primär sowohl anhand des Odds Ratios als auch des Relativen Risikos durchgeführt werden. In begründeten Ausnahmefällen können auch andere Effektmaße zum Einsatz kommen. Bei kategorialen Variablen soll ein geeignetes Effektmaß in Abhängigkeit vom konkreten Endpunkt und den verfügbaren Daten verwendet⁶ werden.

Die Effektschätzer und Konfidenzintervalle aus den Studien sollen mittels Forest Plots zusammenfassend dargestellt werden. Anschließend soll die Einschätzung einer möglichen Heterogenität der Studienergebnisse anhand geeigneter statistische Maße auf Vorliegen von Heterogenität^{7,5} erfolgen. Die Heterogenitätsmaße sind unabhängig von dem Ergebnis der Untersuchung auf Heterogenität immer anzugeben. Ist die Heterogenität der Studienergebnisse nicht bedeutsam (z. B. p-Wert für Heterogenitätsstatistik $\geq 0,05$), soll der gemeinsame (gepoolte) Effekt inklusive Konfidenzintervall dargestellt werden. Bei bedeutsamer Heterogenität sollen die Ergebnisse nur in begründeten Ausnahmefällen gepoolt werden.

⁵ Veroniki AA, Jackson D, Viechtbauer W, Bender R, Knapp G, Kuss O et al. Recommendations for quantifying the uncertainty in the summary intervention effect and estimating the between-study heterogeneity variance in random-effects meta-analysis. *Cochrane Database Syst Rev* 2015: 25-27.

⁶ Deeks JJ, Higgins JPT, Altman DG. Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (Ed). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley; 2008. S. 243-296.

⁷ Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-560.

Außerdem soll untersucht werden, welche Faktoren diese Heterogenität möglicherweise erklären könnten. Dazu zählen methodische Faktoren (siehe Abschnitt 4.2.5.4) und klinische Faktoren, sogenannte Effektmodifikatoren (siehe Abschnitt 4.2.5.5).

Beschreiben Sie die für Meta-Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Nicht zutreffend.

4.2.5.4 Sensitivitätsanalysen

Zur Einschätzung der Robustheit der Ergebnisse sollen Sensitivitätsanalysen hinsichtlich methodischer Faktoren durchgeführt werden. Die methodischen Faktoren bilden sich aus den im Rahmen der Informationsbeschaffung und -bewertung getroffenen Entscheidungen, zum Beispiel die Festlegung von Cut-off-Werten für Erhebungszeitpunkte oder die Wahl des Effektmaßes. Insbesondere die Einstufung des Verzerrungspotenzials der Ergebnisse in die Kategorien „hoch“ und „niedrig“ soll für Sensitivitätsanalysen verwendet werden.

Das Ergebnis der Sensitivitätsanalysen kann die Einschätzung der Aussagekraft der Nachweise beeinflussen.

Begründen Sie die durchgeführten Sensitivitätsanalysen oder den Verzicht auf Sensitivitätsanalysen. Beschreiben Sie die für Sensitivitätsanalysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Nicht zutreffend.

4.2.5.5 Subgruppenmerkmale und andere Effektmodifikatoren

Die Ergebnisse sollen hinsichtlich potenzieller Effektmodifikatoren, das heißt klinischer Faktoren, die die Effekte beeinflussen, untersucht werden. Dies können beispielsweise direkte Patientencharakteristika (Subgruppenmerkmale) sowie Spezifika der Behandlungen (z. B. die Dosis) sein. Im Gegensatz zu den in Abschnitt 4.2.5.4 beschriebenen methodischen Faktoren für Sensitivitätsanalysen besteht hier das Ziel, mögliche Effektunterschiede zwischen Patientengruppen und Behandlungsspezifika aufzudecken. Eine potenzielle Effektmodifikation soll anhand von Homogenitäts- bzw. Interaktionstests oder von Interaktionstermen aus Regressionsanalysen (mit Angabe von entsprechenden Standardfehlern) untersucht werden. Subgruppenanalysen auf der Basis individueller Patientendaten haben in der Regel eine größere Ergebnissicherheit als solche auf Basis von Meta-Regressionsen oder Meta-Analysen unter Kategorisierung der Studien bezüglich der möglichen Effektmodifikatoren, sie sind deshalb zu bevorzugen. Es sollen, soweit sinnvoll, folgende Faktoren bezüglich einer möglichen Effektmodifikation berücksichtigt werden:

- Geschlecht
- Alter
- Krankheitsschwere bzw. –stadium
- Zentrums- und Ländereffekte

Sollten sich aus den verfügbaren Informationen Anzeichen für weitere mögliche Effektmodifikatoren ergeben, können diese ebenfalls begründet einbezogen werden. Die Ergebnisse von in Studien a priori geplanten und im Studienprotokoll festgelegten Subgruppenanalysen für patientenrelevante Endpunkte sind immer darzustellen (zu ergänzenden Kriterien zur Darstellung siehe Abschnitt 4.3.1.3.2).

Bei Identifizierung möglicher Effektmodifikatoren kann gegebenenfalls eine Präzisierung der aus den für die Gesamtgruppe beobachteten Effekten abgeleiteten Aussagen erfolgen. Ergebnisse von Subgruppenanalysen können die Identifizierung von Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen unterstützen.

Benennen Sie die durchgeführten Subgruppenanalysen. Begründen Sie die Wahl von Trennpunkten, wenn quantitative Merkmale kategorisiert werden. Verwenden Sie dabei nach Möglichkeit die in dem jeweiligen Gebiet gebräuchlichen Einteilungen und begründen Sie etwaige Abweichungen. Begründen Sie die durchgeführten Subgruppenanalysen bzw. die Untersuchung von Effektmodifikatoren oder den Verzicht auf solche Analysen. Beschreiben Sie die für diese Analysen eingesetzte Methodik. Begründen Sie, wenn Sie von der oben beschriebenen Methodik abweichen.

Explorative Subgruppenanalysen waren für den primären Endpunkt Anhaltende Remission zu Woche 52 im Statistischen Analyseplan (SAP) der Studie SAPHYR für folgende Faktoren a priori geplant (Tabelle 4-8):

Tabelle 4-8: A priori definierte Subgruppenanalysen der Studie SAPHYR

Subgruppenmerkmal	Ausprägung
Geschlecht	<ul style="list-style-type: none"> • Männlich • Weiblich
Abstammung	<ul style="list-style-type: none"> • Weiß • Andere Abstammung
Geografische Region	<ul style="list-style-type: none"> • Westliche Länder • Südamerika • Rest der Welt
Altersgruppe (Jahre)	<ul style="list-style-type: none"> • < 70 • ≥ 70
Gewicht zu Baseline (kg)	<ul style="list-style-type: none"> • < 60 • ≥ 60 - < 100 • ≥ 100
BMI (kg/m ²)	<ul style="list-style-type: none"> • < 25 • ≥ 25 - < 30 • ≥ 30

CRP zu Baseline	<ul style="list-style-type: none"> • \leq Median • $>$ Median
ESR zu Baseline	<ul style="list-style-type: none"> • \leq Median • $>$ Median
CS-Dosis zu Baseline	<ul style="list-style-type: none"> • \leq Median • $>$ Median
Anzahl der Rezidive vor Screening	<ul style="list-style-type: none"> • 1 • > 1
BMI: Body Mass Index; CRP: C-reaktives Protein; CS: Corticosteroids; ESR: Erythrozyten-Sedimentationsrate. Quelle: Sanofi 2021d	

Begründung für die Trennpunkte bei der Kategorisierung von quantitativen Merkmalen

Alter, Gewicht, BMI und Anzahl der Rezidive vor Screening

Die definierten Trennpunkte orientieren sich an den für die Patientenpopulation und das vorliegende AWG üblichen Bereichen. Alle Einteilungen der dargestellten Trennpunkte wurden gemäß SAP der Studie SAPHYR a priori festgelegt.

Die a priori geplante Subgruppenanalyse zu dem Subgruppenmerkmal „Geografische Region“ wird abweichend vom SAP der Studie SAPHYR dargestellt, da weniger als 10 Patienten mit der Subgruppenausprägung „Südamerika“ identifiziert wurden. Im Rahmen der Nutzenbewertung werden daher die Ausprägungen „Südamerika“ und „Rest der Welt“ zusammengefasst zur neuen Ausprägung „Rest der Welt“.

Die a priori geplanten Subgruppenanalysen zu den subgruppenbildenden Merkmalen „Gewicht zu Baseline“ und „BMI“ wurden nicht durchgeführt, da nicht jede Ausprägung mindestens 10 Patienten umfasste.

Der G-BA bestimmte „MTX in der Kombination mit Glukokortikoiden“ als eine zVT-Option für Sarilumab im vorliegenden AWG (G-BA 2024). Zur Überprüfung eines potenziellen Einflusses von MTX auf die patientenrelevanten Endpunkte wurde das subgruppenbildende Merkmal „MTX Add-on-Therapie“ mit den Subgruppenausprägungen „Ja“ und „Nein“ abweichend vom SAP zu den a priori geplanten Subgruppenanalysen hinzugenommen.

Alle im Dossier dargestellten Subgruppenanalysen für die betrachteten Endpunkte zu Morbidität, gesundheitsbezogener Lebensqualität und Nebenwirkungen erfolgten post-hoc. Die Subgruppenanalysen der Wirksamkeitsendpunkte basierten auf den Daten der ITT-Population. Für die Subgruppenanalysen der Nebenwirkungen wurden die Daten der Sicherheitspopulation herangezogen.

Die Quantifizierung der Unterschiede zwischen den Behandlungen in jeder Subgruppe erfolgte für die binären Endpunkte über die Effektschätzer RR, OR und RD mit den dazugehörigen 95 %-KI und p-Werten mittels logistischer Regression, für die Time-to-Event-Analysen anhand eines Cox-Regressionsmodells über den Effektschätzer HR mit den zugehörigen 95 %-KI und

p-Werten sowie für die kontinuierlichen Endpunkte mittels LS-MWD mit zugehörigen 95 %-KI, p-Werten und Hedges' g mit zugehörigen 95 %-KI.

Mögliche Effektmodifikationen bezüglich der Effektschätzer wurden mit Hilfe des Interaktions-p-Wertes untersucht und tabellarisch dargestellt. Bei Vorliegen einer Interaktion ($p < 0,05$) wurden die Ergebnisse für die jeweiligen Subgruppen separat beschrieben.

Bei der Interpretation der Ergebnisse muss berücksichtigt werden, dass die Durchführung von Interaktionstests mit einer Irrtumswahrscheinlichkeit von 5 % (bei einem p-Wert von 0,05) behaftet ist. Bei der Analyse zahlreicher Subgruppen besteht damit eine hohe Wahrscheinlichkeit, dass die Ergebnisse irgendeiner Subgruppe statistische Signifikanz erreichen, obwohl es sich in Wahrheit um ein zufälliges Ergebnis handelt (IQWiG 2023). Bei einem p-Wert von 0,05, wie er für die Subgruppenanalysen angenommen wurde, würde dies zu einem falsch positiven Signal in einer von 20 Subgruppenanalysen, in denen de facto keine Effektmodifikation vorliegt, führen. Aus diesem Grund und aufgrund des Umstands, dass die betrachtete Studie nicht für Subgruppenanalysen gepowert war, ergibt sich die Notwendigkeit, Aussagen über die Effekte in den Subgruppen vorsichtig und keinesfalls konfirmatorisch zu interpretieren.

4.2.5.6 Indirekte Vergleiche

Zurzeit sind international Methoden in der Entwicklung, um indirekte Vergleiche zu ermöglichen. Es besteht dabei internationaler Konsens, dass Vergleiche einzelner Behandlungsgruppen aus verschiedenen Studien ohne Bezug zu einem gemeinsamen Komparator (häufig als nicht adjustierte indirekte Vergleiche bezeichnet) regelhaft keine valide Analysemethoden darstellen⁸. Eine Ausnahme kann das Vorliegen von dramatischen Effekten sein. An Stelle von nicht adjustierten indirekten Vergleichen sollen je nach Datenlage einfache adjustierte indirekte Vergleiche⁹ oder komplexere Netzwerk-Meta-Analysen (auch als „Mixed Treatment Comparison [MTC] Meta-Analysen“ oder „Multiple Treatment Meta-Analysen“ bezeichnet) für den simultanen Vergleich von mehr als zwei Therapien unter Berücksichtigung sowohl direkter als auch indirekter Vergleiche berechnet werden. Aktuelle Verfahren wurden beispielsweise von Lu und Ades (2004)¹⁰ und Rücker (2012)¹¹ vorgestellt.

Alle Verfahren für indirekte Vergleiche gehen im Prinzip von den gleichen zentralen Annahmen aus. Hierbei handelt es sich um die Annahmen der Ähnlichkeit der eingeschlossenen

⁸ Bender R, Schwenke C, Schmoor C, Hauschke D. Stellenwert von Ergebnissen aus indirekten Vergleichen - Gemeinsame Stellungnahme von IQWiG, GMDS und IBS-DR [online]. [Zugriff: 31.10.2016]. URL: http://www.gmds.de/pdf/publikationen/stellungnahmen/120202_IQWIG_GMDS_IBS_DR.pdf.

⁹ Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50(6): 683-691.

¹⁰ Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23(20): 3105-3124.

¹¹ Rücker G. Network meta-analysis, electrical networks and graph theory. *Res Synth Methods* 2012; 3(4): 312-324.

Studien, der Homogenität der paarweisen Vergleiche und der Konsistenz zwischen direkter und indirekter Evidenz innerhalb des zu analysierenden Netzwerkes. Als Inkonsistenz wird dabei die Diskrepanz zwischen dem Ergebnis eines direkten und eines oder mehreren indirekten Vergleichen verstanden, die nicht mehr nur durch Zufallsfehler oder Heterogenität erklärbar ist¹².

Das Ergebnis eines indirekten Vergleichs kann maßgeblich von der Auswahl des Brückenkomparators bzw. der Brückenkomparatoren abhängen. Als Brückenkomparatoren sind dabei insbesondere Interventionen zu berücksichtigen, für die sowohl zum bewertenden Arzneimittel als auch zur zweckmäßigen Vergleichstherapie mindestens eine direkt vergleichende Studie vorliegt (Brückenkomparatoren ersten Grades).

Insgesamt ist es notwendig, die zugrunde liegende Methodik für alle relevanten Endpunkte genau und reproduzierbar zu beschreiben und die zentralen Annahmen zu untersuchen^{13, 14, 15}

Beschreiben Sie detailliert und vollständig die zugrunde liegende Methodik des indirekten Vergleichs. Dabei sind mindestens folgende Angaben notwendig:

- *Benennung aller potentiellen Brückenkomparatoren ersten Grades und ggf. Begründung für die Auswahl.*
- *Genaue Spezifikation des statistischen Modells inklusive aller Modellannahmen. Bei Verwendung eines Bayes'schen Modells sind dabei auch die angenommenen A-priori-Verteilungen (falls informative Verteilungen verwendet werden, mit Begründung), die Anzahl der Markov-Ketten, die Art der Untersuchung der Konvergenz der Markov-Ketten und deren Startwerte und Länge zu spezifizieren.*
- *Art der Prüfung der Ähnlichkeit der eingeschlossenen Studien.*
- *Art der Prüfung der Homogenität der Ergebnisse direkter paarweiser Vergleiche.*
- *Art der Prüfung der Konsistenzannahme im Netzwerk.*

¹² Schöttker B, Lüthmann D, Boukhemair D, Raspe H. Indirekte Vergleiche von Therapieverfahren. Schriftenreihe Health Technology Assessment Band 88, DIMDI, Köln, 2009.

¹³ Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DJ. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; 338: b1147.

¹⁴ Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study *BMJ* 2011; 343 :d4909

¹⁵ Donegan S, Williamson P, D'Alessandro U, Tudur Smith C. Assessing key assumptions of network meta-analysis: a review of methods. *Res Synth Methods* 2013; 4(4): 291-323.

- *Bilden Sie den Code des Computerprogramms inklusive der einzulesenden Daten in lesbarer Form ab und geben Sie an, welche Software Sie zur Berechnung eingesetzt haben (ggf. inklusive Spezifizierung von Modulen, Prozeduren, Packages etc.; siehe auch Modul 5 zur Ablage des Programmcodes).*
- *Art und Umfang von Sensitivitätsanalysen.*

Nicht zutreffend.

4.3 Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen

In den nachfolgenden Abschnitten sind die Ergebnisse zum medizinischen Nutzen und zum medizinischen Zusatznutzen zu beschreiben. Abschnitt 4.3.1 enthält dabei die Ergebnisse aus randomisierten kontrollierten Studien, die mit dem zu bewertenden Arzneimittel durchgeführt wurden (Evidenzstufen Ia/Ib).

Abschnitt 4.3.2 enthält weitere Unterlagen anderer Evidenzstufen, sofern diese aus Sicht des pharmazeutischen Unternehmers zum Nachweis des Zusatznutzens erforderlich sind. Diese Unterlagen teilen sich wie folgt auf:

- Randomisierte, kontrollierte Studien für einen indirekten Vergleich mit der zweckmäßigen Vergleichstherapie, sofern keine direkten Vergleichsstudien mit der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen (Abschnitt 4.3.2.1)
- Nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2)
- Weitere Untersuchungen (Abschnitt 4.3.2.3)

Falls für die Bewertung des Zusatznutzens mehrere Komparatoren (z.B. Wirkstoffe) herangezogen werden, sind die Aussagen zum Zusatznutzen primär gegenüber der Gesamtheit der gewählten Komparatoren durchzuführen (z. B. basierend auf Meta-Analysen unter gemeinsamer Betrachtung aller direkt vergleichender Studien). Spezifische methodische Argumente, die gegen eine gemeinsame Analyse sprechen (z. B. statistische oder inhaltliche Heterogenität), sind davon unbenommen. Eine zusammenfassende Aussage zum Zusatznutzen gegenüber der zweckmäßigen Vergleichstherapie ist in jedem Fall erforderlich.

4.3.1 Ergebnisse randomisierter kontrollierter Studien mit dem zu bewertenden Arzneimittel

4.3.1.1 Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel

4.3.1.1.1 Studien des pharmazeutischen Unternehmers

Nachfolgend sollen alle Studien (RCT), die an die Zulassungsbehörde übermittelt wurden (Zulassungsstudien), sowie alle Studien (RCT), für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, benannt werden. Beachten Sie dabei folgende Konkretisierungen:

- *Es sollen alle RCT, die der Zulassungsbehörde im Zulassungsdossier übermittelt wurden und deren Studienberichte im Abschnitt 5.3.5 des Zulassungsdossiers enthalten sind, aufgeführt werden. Darüber hinaus sollen alle RCT, für die der pharmazeutische Unternehmer Sponsor ist oder war oder auf andere Weise finanziell beteiligt ist oder war, aufgeführt werden.*

- *Benennen Sie in der nachfolgenden Tabelle nur solche RCT, die ganz oder teilweise innerhalb des in diesem Dokument beschriebenen Anwendungsgebiets durchgeführt wurden. Fügen Sie dabei für jede Studie eine neue Zeile ein.*

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Angabe „Zulassungsstudie ja/nein“, Angabe über die Beteiligung (Sponsor ja/nein), Studienstatus (abgeschlossen, abgebrochen, laufend), Studiendauer, Angabe zu geplanten und durchgeführten Datenschnitten und Therapiearme. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-9: Liste der Studien des pharmazeutischen Unternehmers – RCT mit dem zu bewertenden Arzneimittel

Studie	Zulassungsstudie (ja/nein)	Sponsor (ja/nein)	Status (abgeschlossen / abgebrochen / laufend)	Studiendauer ggf. Datenschnitt	Therapiearme
SAPHYR	ja	ja	abgeschlossen	62 Wochen	Sarilumab 200 mg s.c. q2w + Prednison- Ausschleichen Placebo s.c. q2w + Prednison- Ausschleichen
q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); s.c.: Subkutan (Subcutaneous)					

Geben Sie an, welchen Stand die Information in Tabelle 4-9 hat, d. h. zu welchem Datum der Studienstatus abgebildet wird. Das Datum des Studienstatus soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Informationen in Tabelle 4-9 ist der 22. November 2024.

Geben Sie in der nachfolgenden Tabelle an, welche der in Tabelle 4-9 genannten Studien nicht für die Nutzenbewertung herangezogen wurden. Begründen Sie dabei jeweils die Nichtberücksichtigung. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-10: Studien des pharmazeutischen Unternehmers, die nicht für die Nutzenbewertung herangezogen wurden – RCT mit dem zu bewertenden Arzneimittel

Studienbezeichnung	Begründung für die Nichtberücksichtigung der Studie
Nicht zutreffend	

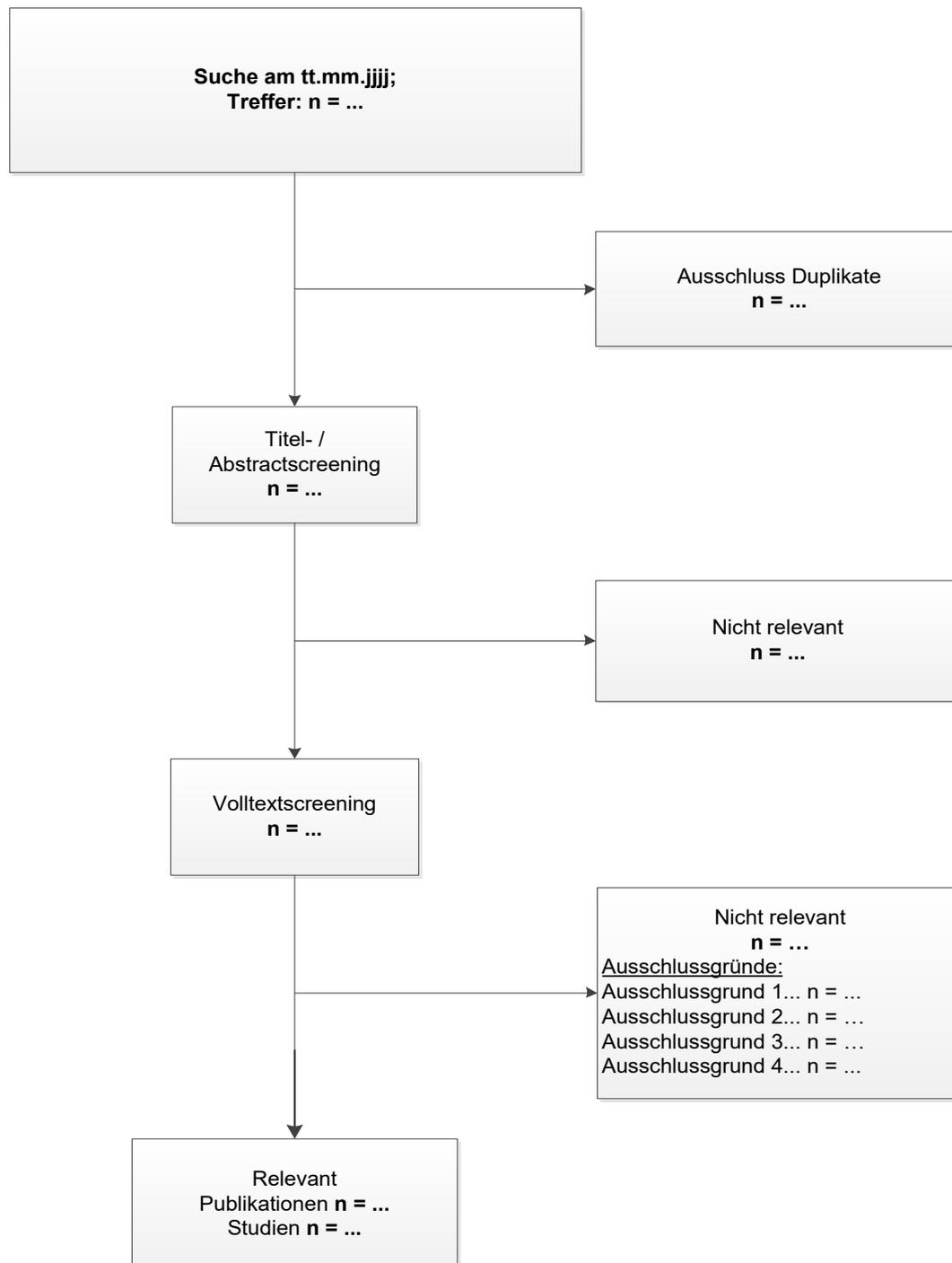
4.3.1.1.2 Studien aus der bibliografischen Literaturrecherche

Beschreiben Sie nachfolgend das Ergebnis der bibliografischen Literaturrecherche. Illustrieren Sie den Selektionsprozess und das Ergebnis der Selektion mit einem Flussdiagramm. Geben Sie dabei an, wie viele Treffer sich insgesamt (d. h. über alle durchsuchten Datenbanken) aus der bibliografischen Literaturrecherche ergeben haben, wie viele Treffer sich nach Entfernung von Dubletten ergeben haben, wie viele Treffer nach Sichtung von Titel und, sofern vorhanden, Abstract als nicht relevant angesehen wurden, wie viele Treffer im Volltext gesichtet wurden, wie viele der im Volltext gesichteten Treffer nicht relevant waren (mit Angabe der Ausschlussgründe) und wie viele relevante Treffer verblieben. Geben Sie zu den relevanten Treffern an, wie vielen Einzelstudien diese zuzuordnen sind. Listen Sie die im Volltext gesichteten und ausgeschlossenen Dokumente unter Nennung des Ausschlussgrunds in Anhang 4-C.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Geben Sie im Flussdiagramm auch das Datum der Recherche an. Die Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Orientieren Sie sich bei der Erstellung des Flussdiagramms an dem nachfolgenden Beispiel.



Zur Identifizierung relevanter Primärpublikationen für RCT mit dem zu bewertenden Arzneimittel wurde am 22. November 2024 eine systematische bibliografische Literaturrecherche durchgeführt. Die Suche und das Screening wurde nach RCT mit dem zu bewertenden Arzneimittel Sarilumab für einen direkten Vergleich gegen die zVT „Therapie nach ärztlicher Maßgabe unter Berücksichtigung von systemischen Glukokortikoiden und der Kombination von Glukokortikoiden mit MTX“ durchgeführt (Tabelle 4-1). Die bibliografische Literaturrecherche zu RCT mit Sarilumab im AWG ergab 26 Treffer, darunter 9 Duplikate. Von

den verbliebenen 17 Treffern wurden 16 nach Titel-/Abstractscreening begründet ausgeschlossen. Somit konnte 1 relevante Publikation identifiziert werden. Hierbei handelt es sich um die Primärpublikation der Zulassungsstudie SAPHYR (Spiera 2023). Das Flussdiagramm des Selektionsprozess für RCT mit dem zu bewertenden Arzneimittel für einen direkten Vergleich gegen die zVT ist in Abbildung 4-1 angeführt. Das vollständige Suchprofil befindet sich in Anhang 4-A.

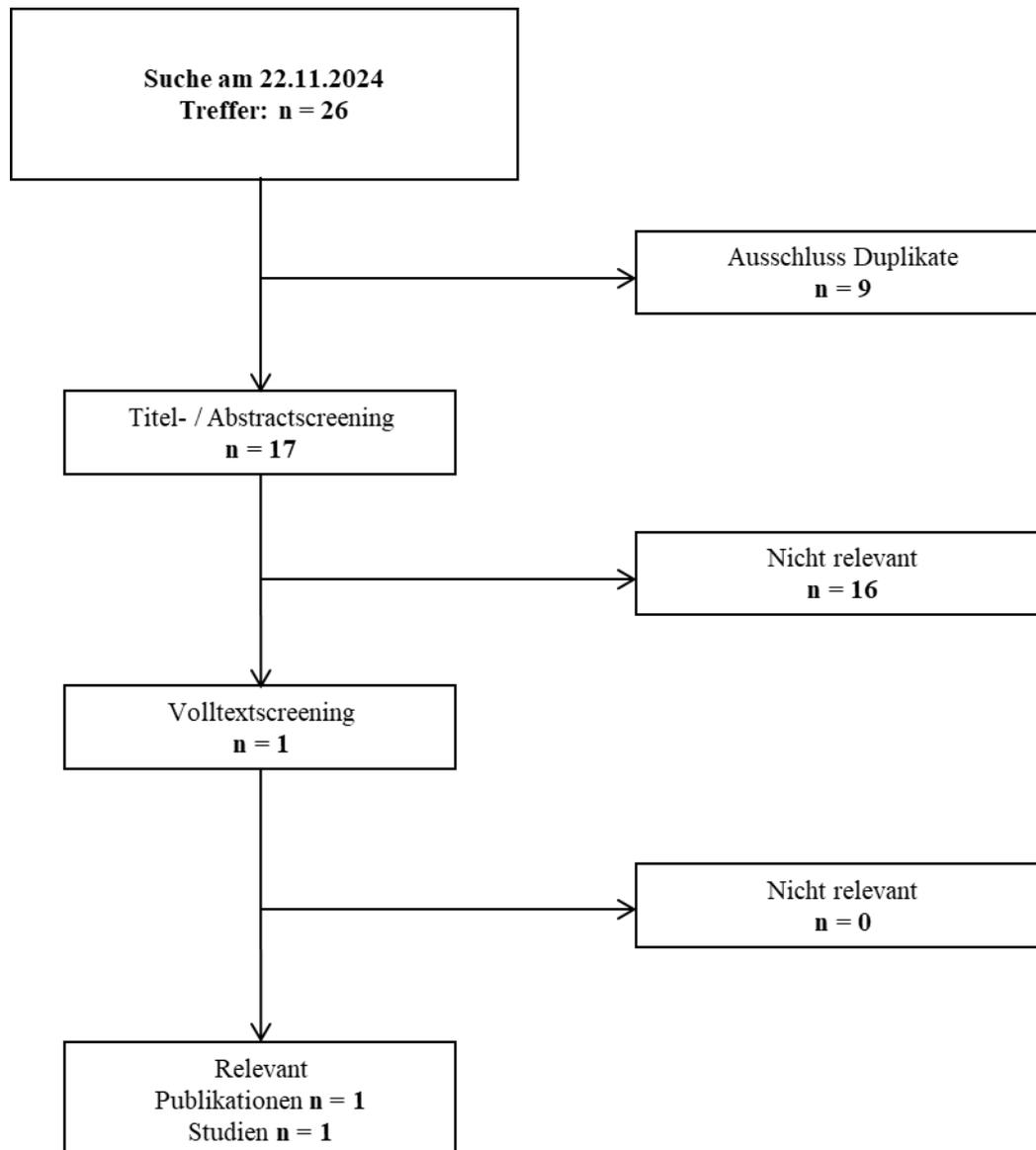


Abbildung 4-1: Flussdiagramm der bibliografischen Literaturrecherche – Suche nach randomisierten, kontrollierten Studien mit dem zu bewertenden Arzneimittel

4.3.1.1.3 Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Suche in Studienregistern/ Studienergebnisdatenbanken identifiziert wurden. Geben Sie dabei an, in

welchem Studienregister / Studienergebnisdatenbank die Studie identifiziert wurde und welche Dokumente dort zur Studie jeweils hinterlegt sind (z. B. Studienregistereintrag, Bericht über Studienergebnisse etc.). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-9) und ob die Studie auch durch die bibliografische Literaturrecherche identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein. Listen Sie die ausgeschlossenen Studien unter Nennung des Ausschlussgrunds in Anhang 4-D.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-11: Relevante Studien (auch laufende Studien) aus der Suche in Studienregistern / Studienergebnisdatenbanken – RCT mit dem zu bewertenden Arzneimittel

Studie	Identifikationsorte (Name des Studienregisters/ der Studienergebnisdatenbank und Angabe der Zitate ^a)	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Status (abgeschlossen/ abgebrochen/ laufend)
SAPHYR NCT03600818 2017-002989-42 U1111-1201-0777	Clinicaltrials.gov (ClinicalTrialsGOV 2022) EU-CTR (EU-CTR 0000) WHO-ICTRP (WHO-ICTRP 2021)	ja	ja	abgeschlossen
<p>a: Zitat des Studienregistereintrags, die Studienregisternummer (NCT-Nummer, EudraCT-Nummer) sowie, falls vorhanden, der im Studienregister/in der Studienergebnisdatenbank aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>EU-CTR: European Union Clinical Trials Register; ICTRP: International Clinical Trials Registry Platform; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); WHO: Weltgesundheitsorganisation (World Health Organization)</p>				

Es wurde eine RCT mit dem zu bewertenden Arzneimittel Sarilumab mit der zVT als Vergleichstherapie im AWG identifiziert. Hierbei handelt es sich um die Zulassungsstudie SAPHYR. Die in dieser Suche gefundenen, aber ausgeschlossenen Studien sind in Anhang 4-D gelistet.

Geben Sie an, welchen Stand die Information in Tabelle 4-11 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dokuments maßgeblichen Zeitpunkt liegen.

Stand der Information in Tabelle 4-11 ist der 22. November 2024.

4.3.1.1.4 Studien aus der Suche auf der Internetseite des G-BA

Beschreiben Sie in der nachfolgenden Tabelle alle relevanten Studien, die durch die Sichtung der Internetseite des G-BA identifiziert wurden. Geben Sie dabei an, welche Dokumente dort hinterlegt sind (z. B. Dossier eines anderen pharmazeutischen Unternehmers, IQWiG Nutzenbewertung). Geben Sie auch an, ob die Studie in der Liste der Studien des pharmazeutischen Unternehmers enthalten ist (siehe Tabelle 4-9) und ob die Studie auch durch die bibliografische Literaturrecherche bzw. Suche in Studienregistern/ Studienergebnisdatenbank identifiziert wurde. Fügen Sie für jede Studie eine neue Zeile ein.

[Anmerkung: „Relevanz“ bezieht sich in diesem Zusammenhang auf die im Abschnitt 4.2.2 genannten Kriterien für den Einschluss von Studien in die Nutzenbewertung.]

Orientieren Sie sich bei Ihren Angaben an der beispielhaften ersten Tabellenzeile.

Tabelle 4-12: Relevante Studien aus der Suche auf der Internetseite des G-BA – RCT mit dem zu bewertenden Arzneimittel

Studie	Relevante Quellen ^a	Studie in Liste der Studien des pharmazeutischen Unternehmers enthalten (ja/nein)	Studie durch bibliografische Literaturrecherche identifiziert (ja/nein)	Studie durch Suche in Studienregistern / Studienergebnisdatenbanken identifiziert (ja/nein)
-	-	-	-	-

a: Quellen aus der Suche auf der Internetseite des G-BA.

Geben Sie an, welchen Stand die Information in Tabelle 4-12 hat, d. h. zu welchem Datum die Recherche durchgeführt wurde. Das Datum der Recherche soll nicht mehr als 3 Monate vor dem für die Einreichung des Dossiers maßgeblichen Zeitpunkt liegen.

Stand der Information in Tabelle 4-12 ist der 22. November 2024. Es konnten keine relevanten Studien zu Sarilumab im vorliegenden AWG auf der Internetseite des G-BA identifiziert werden.

4.3.1.1.5 Resultierender Studienpool: RCT mit dem zu bewertenden Arzneimittel

Benennen Sie in der nachfolgenden Tabelle den aus den verschiedenen Suchschritten (Abschnitte 4.3.1.1.1, 4.3.1.1.2, 4.3.1.1.3 und 4.3.1.1.4) resultierenden Pool relevanter Studien (exklusive laufender Studien) für das zu bewertende Arzneimittel, auch im direkten Vergleich zur zweckmäßigen Vergleichstherapie. Führen Sie außerdem alle relevanten Studien einschließlich aller verfügbaren Quellen in Abschnitt 4.3.1.4 auf. Alle durch die vorhergehenden Schritte identifizierten und in der Tabelle genannten Quellen der relevanten Studien sollen für die Bewertung dieser Studien herangezogen werden.

Folgende Informationen sind in der Tabelle darzulegen: Studienbezeichnung, Studienkategorie und verfügbare Quellen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Hierbei sollen die Studien durch Zwischenzeilenüberschriften ggf. sinnvoll angeordnet werden, beispielsweise nach Therapieschema (Akut-/Langzeitstudien) und jeweils separat nach Art der Kontrolle (Placebo, zweckmäßige Vergleichstherapie, beides). Sollten Sie eine Strukturierung des Studienpools vornehmen, berücksichtigen Sie diese auch in den weiteren Tabellen in Modul 4.

Tabelle 4-13: Studienpool – RCT mit dem zu bewertenden Arzneimittel

Studie	Studienkategorie			verfügbare Quellen ^a		
	Studie zur Zulassung des zu bewertenden Arzneimittels (ja/nein)	gesponserte Studie ^b (ja/nein)	Studie Dritter (ja/nein)	Studienberichte (ja/nein [Zitat])	Registereinträge ^c (ja/nein [Zitat])	Publikation und sonstige Quellen ^d (ja/nein [Zitat])
placebokontrolliert						
Keine ^e						
aktivkontrolliert, zweckmäßige Vergleichstherapie(n)						
SAPHYR ^e	ja	ja	nein	ja (Sanofi 2021a)	ja Clinicaltrials.gov (ClinicalTrialsGOV 2022) EU-CTR (EU-CTR 0000) WHO-ICTRP (WHO-ICTRP 2021)	ja (Spiera 2023)
<p>a: Bei Angabe „ja“ sind jeweils die Zitate der Quelle(n) (z. B. Publikationen, Studienberichte, Studienregistereinträge) mit anzugeben, und zwar als Verweis auf die in Abschnitt 4.6 genannte Referenzliste. Darüber hinaus ist darauf zu achten, dass alle Quellen, auf die in dieser Tabelle verwiesen wird, auch in Abschnitt 4.3.1.4 (Liste der eingeschlossenen Studien) aufgeführt werden.</p> <p>b: Studie, für die der Unternehmer Sponsor war.</p> <p>c: Zitat der Studienregistereinträge sowie, falls vorhanden, der in den Studienregistern aufgelisteten Berichte über Studiendesign und/oder -ergebnisse.</p> <p>d: Sonstige Quellen: Dokumente aus der Suche auf der Internetseite des G-BA.</p> <p>e: Die Studie SAPHYR war placebokontrolliert und trifft die zVT.</p> <p>EU-CTR: European Union Clinical Trials Register; ICTRP: International Clinical Trials Registry Platform; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); WHO: Weltgesundheitsorganisation (World Health Organization).</p>						

4.3.1.2 Charakteristika der in die Bewertung eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

4.3.1.2.1 Studiendesign und Studienpopulationen

Beschreiben Sie das Studiendesign und die Studienpopulation der in die Bewertung eingeschlossenen Studien mindestens mit den Informationen in den folgenden Tabellen. Falls Teilpopulationen berücksichtigt werden, ist die Charakterisierung der Studienpopulation auch für diese Teilpopulation durchzuführen. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Geben Sie bei den Datenschnitten auch den Anlass des Datenschnittes an. Fügen Sie für jede Studie eine neue Zeile ein.

Weitere Informationen zu Studiendesign, Studienmethodik und Studienverlauf sind in Anhang 4-E zu hinterlegen.

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Tabelle 4-14: Charakterisierung der eingeschlossenen Studien – RCT mit dem zu bewertenden Arzneimittel

Studie	Studiendesign	Population	Interventionen (Zahl der randomisierten Patienten)	Studiendauer/ Datenschnitte	Ort und Zeitraum der Durchführung	Primärer Endpunkt; patientenrelevante sekundäre Endpunkte
SAPHYR	Randomisiert, multizentrisch, doppelblind, placebo-kontrolliert, parallel, Phase III	Erwachsene Patienten (≥ 50 Jahre zu Diagnose) mit aktiver PMR, in der Vorgeschichte mindestens 8 Wochen Behandlung mit ≥ 10 mg/Tag Prednison (oder Äquivalent), mindestens ein PMR-Schub bei einer Dosis von ≥ 7,5 mg/Tag Prednison (oder Äquivalent) innerhalb von 12 Wochen vor Screening	Sarilumab (+ Prednison-Ausschleichen) (N=60) Placebo (+ Prednison-Ausschleichen) (N=58)	<u>Screening-Phase:</u> Bis zu 4 Wochen <u>Behandlungsphase:</u> 52 Wochen <u>Nachbeobachtungsphase:</u> 6 Wochen <u>Datenbankschluss:</u> 07. Juli 2021	70 Zentren in 17 Ländern: Australien, Argentinien, Belgien, Deutschland, Estland, Frankreich, Israel, Italien, Japan, Kanada, Niederlande, Russland, Schweiz, Spanien, Ungarn, Vereinigtes Königreich, USA Studienbeginn (FPFV): 09. Oktober 2018 Studienende (LPLV): 19. Mai 2021	<u>Primärer Endpunkt:</u> Anhaltende Remission zu Woche 52 <u>Sekundäre Endpunkte:</u> Einzelkomponenten der anhaltenden Remission, Prednison-Gesamtdosis (oder Äquivalent), Zeit bis zum ersten PMR-Schub, GTI, PMR-AS, VAS (Patienten- und Prüfarztbeurteilung), PRO mittels FACIT-F, EQ-5D, HAQ-DI, SF-36v2, Nebenwirkungen
<p>EQ-5D: European Quality of Life 5 Dimensions; FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; FPFV: Erster Patient erste Visite (First Patient First Visit); GTI: Glukokortikoid-Toxizitätsindex (Glucocorticoid Toxicity Index); HAQ-DI: Health Assessment Questionnaire-Disability Index; LPLV: Letzter Patient letzte Visite (Last Patient Last Visit); N: Anzahl der Patienten; PMR: Polymyalgia rheumatica; PMR-AS: PMR-Aktivitätsscore; PRO: Patientenberichtete Endpunkte (Patient-Reported Outcomes); RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SF-36v2: 36-Item Short-Form Health Survey Version 2; USA: Vereinigte Staaten von Amerika (United States of America).</p> <p>Quelle: Sanofi 2021a</p>						

Tabelle 4-15: Charakterisierung der Interventionen – RCT mit dem zu bewertenden Arzneimittel

Studie	Sarilumab + CST	Placebo + CST	<i>ggf. weitere Spalten mit Behandlungscharakteristika z. B. Vorbehandlung, Behandlung in der Run-in- Phase etc.</i>
SAPHYR	52 Wochen: Sarilumab 200 mg s.c., q2w + Prednison-Ausschleichen: Woche 0 (15 mg/Tag) bis Woche 13 (1 mg/Tag): Prednison oder Kombination aus Prednison und Placebo zu Prednison Woche 14 bis Woche 51: Placebo zu Prednison	52 Wochen: Placebo s.c., q2w + Prednison-Ausschleichen: Woche 0 (15 mg/Tag) bis Woche 51 (1 mg/Tag): Prednison oder Kombination aus Prednison und Placebo zu Prednison	Verbindliche Vorbehandlung mit Prednison (oder Äquivalent): ≥ 10 mg/Tag (mindestens 8 Wochen in der Vorgeschichte) und ≥ 7,5 mg/Tag (innerhalb von 12 Wochen vor Screening) Im Studienverlauf Verabreichung von Prednison (überkapselte Tabletten à 5 mg und 1 mg) bzw. Placebo zu Prednison einmal täglich oral Notwendige Dosiserhöhung bzw. erneute CS-Gabe nach Prednison-Ausschleichen mittels Handelsware
CS: Corticosteroids; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); q2w: Alle zwei Wochen; s.c.: Subkutan (Subcutaneous); RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial). Quelle: Sanofi 2021a			

Tabelle 4-16: Demografische Charakterisierung der Studienpopulation – RCT mit dem zu bewertenden Arzneimittel

Kategorie	Sarilumab + CST (N=60)	Placebo + CST (N=58)
Alter (Jahre) MW (SD) Median (Min : Max) Q1 : Q3	68,8 (7,8) 69,0 (51 : 88) 64,0 : 73,0	69,1 (8,5) 70,0 (52 : 88) 64,0 : 75,0
Altersgruppe (Jahre), n (%) ≥ 50 - < 65 ≥ 65 - < 75 ≥ 75 - < 85 ≥ 85	16 (26,7) 34 (56,7) 9 (15,0) 1 (1,7)	15 (25,9) 26 (44,8) 16 (27,6) 1 (1,7)
Geschlecht, n (%) Männlich Weiblich	15 (25,0) 45 (75,0)	21 (36,2) 37 (63,8)

Kategorie	Sarilumab + CST (N=60)	Placebo + CST (N=58)
Abstammung, n (%)		
Weiß	50 (83,3)	48 (82,8)
Asiatisch	1 (1,7)	2 (3,4)
Nicht berichtet ^a	9 (15,0)	8 (13,8)
Bevölkerungsgruppenzugehörigkeit, n (%)		
Hispanisch oder lateinamerikanisch	6 (10,0)	7 (12,1)
Weder hispanisch noch lateinamerikanisch	45 (75,0)	43 (74,1)
Nicht berichtet ^a	9 (15,0)	8 (13,8)
Geografische Region, n (%)		
Westliche Länder	47 (78,3)	44 (75,9)
Südamerika	5 (8,3)	4 (6,9)
Rest der Welt	8 (13,3)	10 (17,2)
Körpergewicht (kg)		
MW (SD)	74,44 (15,28)	83,49 (16,42)
Median (Min : Max)	73,15 (44,0 : 125,0)	83,50 (57,0 : 127,3)
Q1 : Q3	64,65 : 84,75	70,00 : 94,00
Gewichtsgruppe (kg), n (%)		
< 60	9 (15,0)	4 (6,9)
≥ 60 - < 100	48 (80,0)	43 (74,1)
≥ 100	3 (5,0)	11 (19,0)
BMI (kg/m ²)		
MW (SD)	27,81 (5,42)	29,90 (4,63)
Median (Min : Max)	27,2 (18,6 : 44,8)	30,3 (21,2 : 43,8)
Q1 : Q3	23,75 : 30,79	26,67 : 32,01
BMI-Gruppe (kg/m ²), n (%)		
< 25	18 (30,0)	9 (15,5)
≥ 25 - < 30	24 (40,0)	19 (32,8)
≥ 30	18 (30,0)	30 (51,7)
Raucherstatus, n (%)		
Nie	32 (53,3)	28 (48,3)
Aktiv	11 (18,3)	7 (12,1)
Ehemals	17 (28,3)	23 (39,7)
Alkoholgenuss ^b , n (%)		
Nie	29 (48,3)	25 (43,1)
Gelegentlich	12 (20,0)	14 (24,1)
Mindestens monatlich	5 (8,3)	2 (3,4)
Mindestens wöchentlich	11 (18,3)	12 (20,7)
Mindestens täglich	3 (5,0)	5 (8,6)

Kategorie	Sarilumab + CST (N=60)	Placebo + CST (N=58)
a: Berichtigung von Abstammung und Bevölkerungsgruppenzugehörigkeit in Frankreich nicht erlaubt. b: Innerhalb der letzten 12 Monate. BMI: Body Mass Index; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); Max: Maximum; Min: Minimum; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ausprägung; Q: Quartil; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation). Quelle: Sanofi 2021a		

Tabelle 4-17: Krankheitsspezifische Charakterisierung der Studienpopulation – RCT mit dem zu bewertenden Arzneimittel

Kategorie	Sarilumab + CST (N=60)	Placebo + CST (N=58)
Morgensteifigkeit, n (%)		
Nein	13 (21,7)	13 (22,4)
Ja	47 (78,3)	45 (77,6)
≤ 45 min	12 (25,5)	8 (17,8)
> 45 min	35 (74,5)	37 (82,2)
Morgensteifigkeit (min)		
MW (SD)	85,2 (61,4)	143,0 (215,0)
Median (Min : Max)	60,0 (10 : 270)	90,0 (15 : 1.440)
Q1 : Q3	45,0 : 120,0	60,0 : 180,0
Schulterschmerzen, n (%)		
Ja	43 (71,7)	48 (82,8)
Nein	17 (28,3)	10 (17,2)
Eingeschränkte Schulterbeweglichkeit, n (%)		
Nein	26 (43,3)	21 (36,2)
Ja	34 (56,7)	37 (63,8)
Arm-Anheben über Schulterhöhe	13 (38,2)	12 (32,4)
Arm-Anheben unter Schulterhöhe	8 (23,5)	6 (16,2)
Arm-Anheben bis Schulter	12 (35,3)	19 (51,4)
Kein Arm-Anheben	1 (2,9)	0 (0,0)
Hüftschmerzen, n (%)		
Ja	40 (66,7)	37 (63,8)
Nein	20 (33,3)	21 (36,2)
Eingeschränkte Hüftbeweglichkeit, n (%)		
Ja	24 (40,0)	21 (36,2)
Nein	36 (60,0)	37 (63,8)
Konstitutionelle Symptome, n (%)		
Ja	19 (31,7)	14 (24,1)
Nein	41 (68,3)	44 (75,9)

Kategorie	Sarilumab + CST (N=60)	Placebo + CST (N=58)
Beteiligung anderer Gelenke ^a , n (%)		
Ja	12 (20,0)	6 (10,3)
Nein	48 (80,0)	52 (89,7)
Andere Merkmale, die auf einen PMR-Schub hindeuten, n (%)		
Ja	2 (3,3)	3 (5,2)
Nein	58 (96,7)	55 (94,8)
PMR-Diagnose mittels Ultraschall, n (%)		
Ja	0 (0,0)	0 (0,0)
Nein	29 (48,3)	14 (24,1)
Nicht zutreffend	31 (51,7)	44 (75,9)
Prednison-Dosis (oder Äquivalent) während PMR-Schub (mg)		
MW (SD)	10,53 (3,70)	10,23 (4,16)
Median (Min : Max)	10,00 (5,0 : 20,0)	10,00 (1,0 : 30,0)
Q1 : Q3	7,50 : 12,50	7,50 : 10,00
Patienten mit vorherigen PMR-Schüben, n (%) ^b		
≤ 6 Monate vor Screening	60 (100,0)	58 (100,0)
6 – 12 Monate vor Screening	22 (36,7)	19 (32,8)
> 12 Monate vor Screening	15 (25,0)	18 (31,0)
Anzahl vorheriger PMR-Schübe pro Patient		
MW (SD)	2,8 (2,6)	2,9 (3,2)
Median (Min : Max)	2,0 (1 : 14)	2,0 (1 : 22)
Q1 : Q3	1,0 : 3,0	1,0 : 3,0
Prednison-Dosis (oder Äquivalent) zu Baseline (mg/Tag) ^c		
MW (SD)	12,33 (3,90)	11,31 (3,29)
Median (Min : Max)	11,25 (7,5 : 20,0)	10,00 (7,5 : 20,0)
Q1 : Q3	9,50 : 15,00	8,00 : 15,00
Höchste CS-Dosis ^d für PMR 24 Wochen vor Screening bis Baseline (mg/Tag)		
MW (SD)	20,64 (13,69)	20,36 (20,50)
Median (Min : Max)	15,00 (7,5 : 75,0)	15,00 (7,5 : 150,0)
Q1 : Q3	14,50 : 20,00	10,00 : 20,00
Dauer des CS-Gebrauchs ^e für PMR 24 Wochen vor Screening bis Baseline (Tage)		
MW (SD)	144,0 (36,4)	154,8 (24,5)
Median (Min : Max)	168,0 (41 : 168)	168,0 (78 : 168)
Q1 : Q3	122,0 : 168,0	148,0 : 168,0
Dauer der PMR von Diagnose bis Baseline (Tage)		
n	54	50
MW (SD)	660,0 (818,0)	600,2 (680,5)
Median (Min : Max)	291,5 (78 : 3.992)	309,5 (66 : 2.784)
Q1 : Q3	160,0 : 865,0	144,0 : 883,0

Kategorie	Sarilumab + CST (N=60)	Placebo + CST (N=58)
CRP zu Baseline (mg/l) MW (SD) Median (Min : Max) Q1 : Q3	9,97 (9,79) 6,76 (0,5 : 38,2) 1,99 : 15,10	8,92 (11,20) 5,69 (0,1 : 62,3) 2,24 : 9,45
ESR zu Baseline (mm/h) MW (SD) Median (Min : Max) Q1 : Q3	31,15 (24,55) 25,00 (2,0 : 115,0) 12,50 : 40,00	26,93 (17,09) 22,00 (5,0 : 85,0) 15,00 : 35,00
VAS – Krankheitsaktivität, Prüfarztbeurteilung n MW (SD) Median (Min : Max) Q1 : Q3	49 52,9 (23,5) 60,0 (0 : 98) 40,0 : 69,0	52 53,6 (25,1) 61,5 (0 : 89) 36,0 : 72,5
GTI zu Baseline (0-538) MW (SD) Median (Min : Max) Q1 : Q3	100,8 (42,9) 97,0 (0 : 228) 76,0 : 121,5	115,3 (51,0) 112,5 (0 : 267) 86,0 : 141,0
Vorherige biologische DMARD, n (%) Adalimumab Tocilizumab	1 (1,7) 1 (1,7) 0 (0,0)	1 (1,7) 0 (0,0) 1 (1,7)
Vorherige nicht-biologische DMARD, n (%) Methotrexat Methotrexat-Natrium Leflunomid Azathioprin Hydroxychloroquin Hydroxychloroquin-Sulfat	6 (10,0) 2 (3,3) 3 (5,0) 2 (3,3) 0 (0,0) 0 (0,0) 1 (1,7)	11 (19,0) 9 (15,5) 1 (1,7) 1 (1,7) 1 (1,7) 1 (1,7) 0 (0,0)
<p>a: Andere entzündete Gelenke (z. B. Hände, Handgelenke, Knie etc.) mit Bezug zur Krankheitsaktivität der PMR. b: Mehrere vorherige Schübe pro Patient und damit Zählung in mehr als einer Kategorie möglich. c: Basierend auf der letzten Dosis vor der Baseline-Visite. d: Einschließlich Prednison oder Äquivalent. e: Gesamtdauer des CS-Gebrauchs ohne Dosierungsunterbrechungen.</p> <p>CRP: C-reaktives Protein; CS: Corticosteroids; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); DMARD: Krankheitsmodifizierende Antirheumatika (Disease-Modifying Anti-Rheumatic Drugs); ESR: Erythrozyten-Sedimentationsrate; GTI: Glukokortikoid-Toxizitätsindex; Max: Maximum; Min: Minimum; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ausprägung; PMR: Polymyalgia rheumatica; Q: Quartil; RCT: Randomisierte, kontrollierte Studie Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation); VAS: Visuelle Analogskala.</p> <p>Quelle: Sanofi 2021a</p>		

Beschreiben Sie die Studien zusammenfassend. In der Beschreibung der Studien sollten Informationen zur Behandlungsdauer sowie zu geplanter und tatsächlicher Beobachtungsdauer enthalten sein. Sofern sich die Beobachtungsdauer zwischen den relevanten Endpunkten unterscheidet, sind diese unterschiedlichen Beobachtungsdauern endpunktbezogen anzugeben.

Beschreiben Sie zudem, ob und aus welchem Anlass verschiedene Datenschnitte durchgeführt wurden oder noch geplant sind. Geben Sie dabei auch an, ob diese Datenschnitte jeweils vorab (d.h. im statistischen Analyseplan) geplant waren. In der Regel ist nur die Darstellung von a priori geplanten oder von Zulassungsbehörden geforderten Datenschnitten erforderlich. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Sollte es Unterschiede zwischen den Studien geben, weisen Sie in einem erläuternden Text darauf hin.

Studie SAPHYR

Bei der Studie SAPHYR handelt es sich um eine abgeschlossene multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte Phase-III-Studie an erwachsenen Patienten (Alter ≥ 50 Jahre zu Diagnose) mit aktiver PMR, die in der Vorgeschichte mindestens 8 Wochen mit ≥ 10 mg/Tag Prednison (oder Äquivalent) behandelt wurden und innerhalb von 12 Wochen vor Screening mindestens einen eindeutigen PMR-Schub bei einer Dosis von $\geq 7,5$ mg/Tag Prednison (oder Äquivalent) erlitten. Während der Screening-Phase mussten die Patienten mit Prednison in einer Dosis von $\geq 7,5$ mg/Tag (oder Äquivalent) und ≤ 20 mg/Tag therapiert sein. Die Patienten wurden im Verhältnis 1:1 entweder Sarilumab 200 mg q2w oder Placebo zu Sarilumab q2w zugeteilt. Zusätzlich erhielten alle Patienten Prednison, das im Sarilumab-Arm von 15 mg/Tag in den ersten beiden Wochen schrittweise auf 1 mg/Tag in Woche 13 ausgeschlichen wurde. Im Kontrollarm wurde Prednison von 15 mg/Tag in den ersten beiden Wochen schrittweise auf 1 mg/Tag in Woche 51 ausgeschlichen. In beiden Studienarmen erhielten die Patienten im gesamten Behandlungsverlauf zur Aufrechterhaltung der Verblindung entweder Prednison, eine Kombination aus Prednison und Placebo zu Prednison oder ausschließlich Placebo zu Prednison. Anstelle der verblindeten Gabe von Prednison bzw. Placebo zu Prednison konnten kommerzielle CS, einschließlich Prednison, in beiden Studienarmen von Woche 12 bis Woche 51 als Notfallmedikation verabreicht werden. Patienten, bei denen die PMR trotz Gabe kommerzieller CS symptomatisch blieb, konnten mit anderen Wirkstoffen weiterbehandelt werden, z. B. nicht-biologischen Immunsuppressiva.

Für die vorliegende Nutzenbewertung basiert die Analyse der patientenrelevanten Wirksamkeitsendpunkte auf der ITT-Population, die alle randomisierten Patienten umfasste. Aufgrund der Rekrutierungsproblematik während der COVID-19-Pandemie wurde die Fallzahl von ursprünglich geplanten 280 auf 118 randomisierte Patienten reduziert. Die ITT-Population enthält somit 118 Patienten, 60 im Sarilumab-Arm und 58 im Kontrollarm. Die demografischen Patientencharakteristika zu Baseline fielen für die beiden Studienarme vergleichbar aus: das mediane Alter lag bei 69,0 Jahren im Sarilumab-Arm bzw. 70,0 Jahren im Kontrollarm, die Mehrheit der Patienten war weiblich (75,0 % bzw. 63,8 %), hatte eine weiße Hautfarbe (83,3 % bzw. 82,8 %) und kam aus einem westlichen Land (78,3 % bzw. 75,9 %). Jeweils die Hälfte der Patienten hatte nie geraucht (53,3 % bzw. 48,3 %) und in den letzten 12 Monaten keinen Alkohol konsumiert (48,3 % bzw. 43,1 %) (Tabelle 4-16). Die krankheitsspezifischen Patientencharakteristika bestätigen die Vergleichbarkeit beider Studienarme: Morgensteifigkeit beklagten 78,3 % bzw. 77,6 %, Schulterschmerzen 71,7 % bzw. 82,8 % und Hüftschmerzen

66,7 % bzw. 63,8 % der Patienten. Symptome in anderen Gelenken mit Bezug zur PMR-Aktivität wurden von lediglich 20,0 % bzw. 10,3 % der Patienten berichtet. Die Prüfarzte beurteilten die Krankheitsaktivität mittels VAS mit medianen 60,0 bzw. 61,5 Punkten. In beiden Studienarmen lag die Anzahl vorheriger PMR-Schübe im Median bei 2,0 und der Median der Prednison-Dosis (oder Äquivalent) während eines Schubs bei 10,0 mg. Die Zeitspanne zwischen PMR-Diagnose und Baseline betrug im Median 291,5 bzw. 309,5 Tage. Bis zum Tag vor Baseline erhielten die Patienten im Median 11,25 mg bzw. 10,00 mg Prednison (oder Äquivalent) pro Tag. In beiden Studienarmen waren jeweils 1,7 % der Patienten mit biologischen DMARD und 10,0 % bzw. 19,0 % mit nicht-biologischen DMARD vortherapiert (Tabelle 4-17).

Die Studiendauer pro Patient betrug etwa 62 Wochen, bestehend aus einer bis zu 4-wöchigen Screening-Phase, einer 52-wöchigen Behandlungsphase und einer 6-wöchigen Nachbeobachtungsphase.

Primärer Wirksamkeitsendpunkt der Studie war der Anteil der Patienten mit anhaltender Remission zu Woche 52. Zu den sekundären Endpunkten zählten unter anderem die kumulative Corticosteroid-Gesamtdosis während der 52-wöchigen Behandlungsphase, die Einzelkomponenten der anhaltenden Remission zu Woche 52 (Erreichen der Remission spätestens zu Woche 12, kein Krankheitsschub von Woche 12 bis Woche 52, anhaltende Reduktion des CRP von Woche 12 bis Woche 52 sowie erfolgreiches Prednison-Ausschleichen von Woche 12 bis Woche 52) und Nebenwirkungen. Neben den sekundären Endpunkten wurde die Lebensqualität als patientenrelevanter Endpunkt erfasst.

Es waren keine Zwischenanalysen geplant. Die letzte Visite des letzten Patienten fand am 19. Mai 2021 statt. Die Datenbank wurde mit Datum 07. Juli 2021 geschlossen. Die bis zu diesem Zeitpunkt erhobenen Daten werden für die Darstellung der Endpunktergebnisse in Abschnitt 4.3.1.3 herangezogen.

Die mediane Behandlungsdauer lag in beiden Studienarmen bei 363,0 Tagen. Ein Patient, der in den Sarilumab-Arm randomisiert wurde, erhielt keine Studienmedikation. Von den übrigen 59 Patienten beendeten 42 (70,0 %) die Behandlung protokollgerecht, während 17 Patienten (28,3 %) sie frühzeitig beendeten. Im Kontrollarm beendeten 36 Patienten (62,1 %) die Behandlung protokollgerecht, 22 (37,9 %) brachen sie frühzeitig ab. Als häufigster Grund für den Behandlungsabbruch wurden im Sarilumab-Arm unerwünschte Ereignisse bei 11,7 % der Patienten, im Kontrollarm mangelnde Wirksamkeit bei 15,5 % der Patienten angegeben (Sanofi 2021a).

Eine anhaltende Remission zu Woche 52 erreichten 17 Patienten (28,3 %) im Sarilumab-Arm und 6 Patienten (10,3 %) im Kontrollarm. Eine Remission spätestens zu Woche 12 zeigte sich bei 28 (46,7 %) bzw. 22 Patienten (37,9 %), ohne PMR-Schub von Woche 12 bis Woche 52 blieben 33 (55,0 %) bzw. 19 Patienten (32,8 %), eine dauerhafte Reduktion des CRP von Woche 12 bis Woche 52 wurde für 40 (66,7 %) bzw. 26 Patienten (44,8 %) gemessen, ferner konnten 30 (50,0 %) bzw. 14 Patienten (24,1 %) Prednison von Woche 12 bis Woche 52 ohne Bedarf an CS-Notfallmedikation erfolgreich ausschleichen (Sanofi 2021a).

Die Sicherheitspopulation, die alle Patienten berücksichtigte, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten, umfasste 117 Patienten, 59 im Sarilumab-Arm und 58 im Kontrollarm. Der Anteil der Patienten mit mindestens einem UE betrug 94,9 % im Sarilumab-Arm und 84,5 % im Kontrollarm. Todesfälle traten in der Studie SAPHYR nicht auf (Sanofi 2021a).

Die Beurteilung des Gesundheitsstatus durch den Patienten erfolgte mittels FACIT-F, EQ-5D VAS und HAQ-DI. Zur patientenberichteten gesundheitsbezogenen Lebensqualität kam die Version 2 des SF-36 zum Einsatz. Die Ergebnisse der Fragebögen zum Gesundheitsstatus fielen für beide Studienarme vergleichbar aus (Sanofi 2021a).

Übertragbarkeit auf den deutschen Versorgungskontext

In der Studie SAPHYR wurden insgesamt 77,1 % der Patienten in westlichen Ländern rekrutiert und behandelt. Das mediane Alter der Gesamtpopulation zu Studienbeginn lag bei 69,5 Jahren, 69,5 % der Patienten waren Frauen. Dies steht in Übereinstimmung mit der deutsch/österreichisch/schweizerischen S3-Leitlinie zur Behandlung der Polymyalgia rheumatica, die den Krankheitsbeginn fast ausschließlich bei Personen oberhalb des 50. Lebensjahres angibt, mit einem Verhältnis von betroffenen Frauen zu Männern von 3:1 (Buttgereit 2018). Im Rahmen der Studie erfolgte die Verabreichung von Sarilumab darüber hinaus entsprechend der Zulassung (Sanofi 2024a). Es ist somit davon auszugehen, dass die Studienpopulation den deutschen Versorgungskontext abbildet.

4.3.1.2.2 Verzerrungspotenzial auf Studienebene

Bewerten Sie das Verzerrungspotenzial der RCT auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-18: Verzerrungspotenzial auf Studienebene – RCT mit dem zu bewertenden Arzneimittel

Studie	Adäquate Erzeugung der Randomisierungssequenz	Verdeckung der Gruppenzuteilung	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial auf Studienebene
			Patient	Behandelnde Personen			
SAPHYR	ja	ja	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Randomisierung und Gruppenzuteilung erfolgten in der Studie SAPHYR adäquat und verdeckt mit Hilfe einer interaktiven Dialogtechnologie (Interactive Response Technology, IRT). Sowohl Patienten als auch behandelnde Personen waren hinsichtlich der Zuteilung zu Sarilumab oder Placebo zu Sarilumab verblindet. Die zusätzliche Prednison-Gabe geschah ebenfalls doppelblind. Im Sarilumab-Arm wurde Prednison von Woche 2 bis einschließlich Woche 13 schrittweise ausgeschlichen. Im Kontrollarm wurde Prednison von Woche 2 bis einschließlich Woche 51 schrittweise ausgeschlichen. Zur Aufrechterhaltung der Verblindung erhielten die Patienten im gesamten Behandlungsverlauf entweder Prednison oder eine Kombination aus Prednison oder Placebo zu Prednison. Anstelle der verblindeten Gabe von Prednison bzw. Placebo zu Prednison konnten kommerzielle Corticosteroide, einschließlich Prednison, in beiden Studienarmen von Woche 12 bis Woche 51 als Notfallmedikation verabreicht werden. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da die Analysen prospektiv im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Darüber hinaus lagen keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial auf Studienebene als niedrig eingestuft (Tabelle 4-18).

4.3.1.3 Ergebnisse aus randomisierten kontrollierten Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens beruht. Geben Sie dabei an, welche dieser Endpunkte in den relevanten Studien jeweils untersucht wurden. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Tabellenzeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-19: Matrix der Endpunkte in den eingeschlossenen RCT mit dem zu bewertenden Arzneimittel

Studie	Mortalität	Morbidität	Gesundheits- bezogene Lebensqualität	Nebenwirkungen
SAPHYR	Gesamtmortalität ^a	Anhaltende Remission zu Woche 52, inkl. Einzelkomponenten Kumulative CS-Gesamtdosis Zeit bis zum ersten PMR-Schub PMR-AS, inkl. Einzelkomponenten und VAS – Krankheitsaktivität, Prüfarztbeurteilung FACIT-F EQ-5D VAS HAQ-DI	SF-36v2	UE gesamt Schwere UE SUE Therapieabbruch aufgrund von UE UE von besonderem Interesse
a: Beurteilung anhand von UE mit Todesfolge.				

Studie	Mortalität	Morbidität	Gesundheits- bezogene Lebensqualität	Nebenwirkungen
CS: Corticosteroids; EQ-5D: European Quality of Life 5 Dimensions; FACIT--F: Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI: Health Assessment Questionnaire-Disability Index; PMR: Polymyalgia rheumatica; PMR-AS: PMR-Aktivitätsscore; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SF-36v2: 36-Item Short-Form Health Survey Version 2; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala.				

4.3.1.3.1 Endpunkt – RCT

Die Ergebnisdarstellung für jeden Endpunkt umfasst 3 Abschnitte. Zunächst soll für jede Studie das Verzerrungspotenzial auf Endpunktebene in einer Tabelle zusammengefasst werden. Dann sollen die Ergebnisse der einzelnen Studien zu dem Endpunkt tabellarisch dargestellt und in einem Text zusammenfassend beschrieben werden. Anschließend sollen die Ergebnisse, wenn möglich und sinnvoll, in einer Meta-Analyse zusammengefasst und beschrieben werden.

Die tabellarische Darstellung der Ergebnisse für den jeweiligen Endpunkt soll mindestens die folgenden Angaben enthalten:

- Ergebnisse der ITT-Analyse
- Zahl der Patienten, die in die Analyse eingegangen sind inkl. Angaben zur Häufigkeit von und zum Umgang mit nicht oder nicht vollständig beobachteten Patienten (bei Verlaufsbeobachtungen pro Messzeitpunkt)
- dem Endpunkt entsprechende Kennzahlen pro Behandlungsgruppe
- bei Verlaufsbeobachtungen Werte zu Studienbeginn und Studienende inklusive Standardabweichung
- bei dichotomen Endpunkten die Anzahlen und Anteile pro Gruppe sowie Angabe des relativen Risikos, des Odds Ratios und der absoluten Risikoreduktion
- entsprechende Maße bei weiteren Messniveaus
- Effektschätzer mit zugehörigem Standardfehler
- Angabe der verwendeten statistischen Methodik inklusive der Angabe der Faktoren, nach denen ggf. adjustiert wurde.

Unterschiedliche Beobachtungszeiten zwischen den Behandlungsgruppen sollen durch adäquate Analysen (z.B. Überlebenszeitanalysen) adressiert werden, und zwar für alle Endpunkte (einschließlich UE nach den nachfolgend genannten Kriterien), für die eine solche Analyse aufgrund deutlich unterschiedlicher Beobachtungszeiten erforderlich ist.

Bei Überlebenszeitanalysen soll die Kaplan-Meier-Kurve einschließlich Angaben zu den Patienten unter Risiko im Zeitverlauf (zu mehreren Zeitpunkten) abgebildet werden. Dabei ist für jeden Endpunkt, für den eine solche Analyse durchgeführt wird, eine separate Kaplan-Meier-Kurve darzustellen.

Zu mit Skalen erhobenen patientenberichteten Endpunkten (z.B. zur gesundheitsbezogenen Lebensqualität oder zu Symptomen) sind immer auch die Werte im Studienverlauf anzugeben,

auch als grafische Darstellung, sowie eine Auswertung, die die über den Studienverlauf ermittelten Informationen vollständig berücksichtigt (z.B. als Symptomlast über die Zeit, geschätzt mittels MMRM-Analyse [falls aufgrund der Datenlage geeignet]). Die Auswertung von Responderanalysen mittels klinischer Relevanzschwellen bei komplexen Skalen soll nach dem folgenden Vorgehen erfolgen:

1. Falls in einer Studie Responderanalysen unter Verwendung einer MID präspezifiziert sind und das Responsekriterium mindestens 15 % der Skalenspannweite des verwendeten Erhebungsinstruments entspricht, sind diese Responderanalysen für die Bewertung darzustellen.

2. Falls präspezifiziert Responsekriterien im Sinne einer MID unterhalb von 15 % der Skalenspannweite liegen, bestehen in diesen Fällen und solchen, in denen gar keine Responsekriterien präspezifiziert wurden, aber stattdessen Analysen kontinuierlicher Daten zur Verfügung stehen, verschiedene Möglichkeiten. Entweder können post hoc spezifizierte Analysen mit einem Responsekriterium von genau 15 % der Skalenspannweite dargestellt werden. Alternativ können Analysen der kontinuierlichen Daten dargestellt werden, für die Relevanzbewertung ist dabei auf ein allgemeines statistisches Maß in Form von standardisierten Mittelwertdifferenzen (SMDs, in Form von Hedges' g) zurückzugreifen. Dabei ist eine Irrelevanzschwelle als Intervall von -0,2 bis 0,2 zu verwenden: Liegt das zum Effektschätzer korrespondierende Konfidenzintervall vollständig außerhalb dieses Irrelevanzbereichs, wird davon ausgegangen, dass die Effektstärke nicht in einem sicher irrelevanten Bereich liegt. Dies soll gewährleisten, dass der Effekt hinreichend sicher mindestens als klein angesehen werden kann.

3. Liegen sowohl geeignete Responderanalysen (Responsekriterium präspezifiziert mindestens 15 % der Skalenspannweite oder post hoc genau 15 % der Skalenspannweite) als auch Analysen stetiger Daten vor, sind die Responderanalysen darzustellen.

Zu unerwünschten Ereignissen (UE) sind folgende Auswertungen vorzulegen:

1. Gesamtrate UE,
2. Gesamtrate schwerwiegender UE (SUE),
3. Gesamtrate der Abbrüche wegen UE,
4. Gesamtraten von UE differenziert nach Schweregrad, sofern dies in der/den relevante/n Studie/n erhoben wurde (z.B. gemäß CTCAE und/oder einer anderen etablierten bzw. validierten indikationsspezifischen Klassifikation) einschließlich einer Abgrenzung schwerer und nicht schwerer UE,
5. zu den unter 1, 2 und 4 genannten Kategorien (UE ohne weitere Differenzierung, SUE, UE differenziert nach Schweregrad) soll zusätzlich zu den Gesamtraten die Darstellung nach Organsystemen und Einzelereignissen (als System Organ Class [SOCs] und Preferred Terms [PT] nach MedDRA) jeweils nach folgenden Kriterien erfolgen:

- UE (unabhängig vom Schweregrad): Ereignisse, die bei mindestens 10 % der Patienten in einem Studienarm aufgetreten sind
- Schwere UE (z.B. CTCAE-Grad ≥ 3) und SUE: Ereignisse, die bei mindestens 5% der Patienten in einem Studienarm aufgetreten sind
- zusätzlich für alle Ereignisse unabhängig vom Schweregrad: Ereignisse, die bei mindestens 10 Patienten UND bei mindestens 1 % der Patienten in einem Studienarm aufgetreten sind.

6. A priori definierte UE von besonderem Interesse [AESI]) sowie prädefinierte SOC-übergreifende UE-Auswertungen (z.B. als Standardised MedDRA Queries, SMQs) sollen unabhängig von der Ereignisrate dargestellt werden und zwar differenziert nach Schweregrad (dargestellt als Gesamtrate und differenziert nach Schweregrad, nicht schwer, schwer, schwerwiegend).

7. zu Kategorie 3: Die Abbruchgründe auf SOC/PT-Ebene müssen vollständig, jedoch nur deskriptiv dargestellt werden.

Sofern bei der Erhebung unerwünschter Ereignisse erkrankungsbezogenen Ereignisse (z. B. Progression, Exazerbation) berücksichtigt werden (diese Ereignisse also in die UE-Erhebung eingehen), sollen für die Gesamtraten (UE, schwere UE und SUE) zusätzliche UE-Analysen durchgeführt werden, bei denen diese Ereignisse unberücksichtigt bleiben. Alle Auswertungen zu UE können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine Darstellung ausschließlich in Modul 5 ist nicht ausreichend. Davon unbenommen sind die Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE), sowie die für die Gesamtaussage zum Zusatznutzen herangezogenen Ergebnisse im vorliegenden Abschnitt darzustellen.

Auswertungen zu den im Abschnitt 4.3.1.2.1 aufgeführten Datenschnitten sollen vollständig, d.h. für alle erhobenen relevanten Endpunkte, durchgeführt und vorgelegt werden. Das gilt auch dann wenn ein Datenschnitt ursprünglich nur zur Auswertung einzelner Endpunkte geplant war. Auf die Darstellung der Ergebnisse einzelner Endpunkte eines Datenschnitts bzw. eines gesamten Datenschnitts kann verzichtet werden, wenn hierdurch kein wesentlicher Informationsgewinn gegenüber einem anderen Datenschnitt zu erwarten ist (z. B. wenn die Nachbeobachtung zu einem Endpunkt bereits zum vorhergehenden Datenschnitt nahezu vollständig war oder ein Datenschnitt in unmittelbarer zeitlicher Nähe zu einem anderen Datenschnitt liegt).

Falls für die Auswertung eine andere Population als die ITT-Population herangezogen wird, soll diese benannt (z.B. Safety-Population) und definiert werden.

Sofern mehrere Studien vorliegen, sollen diese in einer Meta-Analyse zusammengefasst werden, wenn die Studien aus medizinischen (z. B. Patientengruppen) und methodischen (z. B. Studiendesign) Gründen ausreichend vergleichbar sind. Es ist jeweils zu begründen, warum

eine Meta-Analyse durchgeführt wurde oder warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Sofern die vorliegenden Studien für eine Meta-Analyse geeignet sind, sollen die Meta-Analysen als Forest-Plot dargestellt werden. Die Darstellung soll ausreichende Informationen zur Einschätzung der Heterogenität der Ergebnisse zwischen den Studien in Form von geeigneten statistischen Maßzahlen enthalten (siehe Abschnitt 4.2.5.3). Eine Gesamtanalyse aller Patienten aus mehreren Studien ohne Berücksichtigung der Studienzugehörigkeit (z. B. Gesamt-Vierfeldertafel per Addition der Einzel-Vierfeldertafeln) soll vermieden werden, da so die Heterogenität nicht eingeschätzt werden kann.

4.3.1.3.1.1 Endpunkt Gesamtmortalität - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-20: Operationalisierung des Endpunkts Gesamtmortalität

Studie	Operationalisierung
SAPHYR	Die Beurteilung der Gesamtmortalität erfolgte anhand der UE mit Todesfolge, die von der Unterzeichnung der Einwilligungserklärung bis zum Studienende hinaus erfasst wurden. SUE oder UE von besonderem Interesse, die zum Studienende nicht behoben waren, wurden bis zur Normalisierung, Stabilisierung oder bis zum Tod weiter beobachtet. Die Analyse basierte auf den Daten der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten.
UE: Unerwünschtes Ereignis; SUE: Schwerwiegendes unerwünschtes Ereignis.	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-21: Bewertung des Verzerrungspotenzials für den Endpunkt Gesamtmortalität in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Gesamtmortalität ist als niedrig einzustufen, da eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Gesamtmortalität als niedrig eingestuft (Tabelle 4-21).

Stellen Sie die Ergebnisse für den Endpunkt Gesamtmortalität für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-22: Ergebnisse für Gesamtmortalität aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Endpunkt	Sarilumab + CST (N=60)	Placebo + CST (N=58)
Todesfälle, n (%)	0 (0,0)	0 (0,0)
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial). Quelle: Sanofi 2021a		

In der Studie SAPHYR traten während des gesamten Studienverlaufs keine Todesfälle auf.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt Gesamtmortalität auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus RCT mit dem zu bewertenden Arzneimittel fortlaufend in einem eigenen Abschnitt dar.

4.3.1.3.1.2 Endpunkt Anhaltende Remission zu Woche 52 - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-23: Operationalisierung des Endpunkts Anhaltende Remission zu Woche 52

Studie	Operationalisierung
SAPHYR	<p>Die anhaltende Remission zu Woche 52 setzte sich zusammen aus den vier Einzelkomponenten:</p> <ol style="list-style-type: none"> 1. Remission spätestens zu Woche 12 2. Kein Krankheitsschub von Woche 12 bis 52 3. Anhaltende CRP-Reduktion von Woche 12 bis 52 4. Erfolgreiches Prednison-Ausschleichen von Woche 12 bis 52 <p>Ad 1. Remission wurde definiert als Auflösung der Zeichen und Symptome der PMR und CRP-Normalisierung.</p> <p>Die Beurteilung der Zeichen und Symptome der PMR beinhaltet unter anderem:</p> <ul style="list-style-type: none"> ○ Morgensteifigkeit und/oder Schmerzen in Nacken, Schulter und/oder Beckengürtel ○ Eingeschränkte Bewegungsfähigkeit in Schultern und/oder Beckengürtel ○ Konstitutionelle Symptome wie Fatigue, Gewichtsverlust und leichtes Fieber ○ Andere Auffälligkeiten, die nach Ansicht des Prüfarztes auf einen PMR-Schub hinweisen <p>Ad 2. Krankheitsschub wurde definiert als entweder das Wiederauftreten von Zeichen und Symptomen, die eine Erhöhung der CS-Dosis erforderlich machten, oder als Anstieg der ESR, der auf eine aktive PMR zurückzuführen ist und eine Erhöhung der CS-Dosis erforderlich machte. Erhöhung der CS-Dosis wurde definiert als:</p> <ul style="list-style-type: none"> ○ Jede Dosiserhöhung während des protokollgerechten Prednison-Ausschleichens ○ Erneute CS-Gabe nach Beendigung des protokollgerechten Ausschleichens <p>Ad 3. Anhaltende CRP-Reduktion wurde definiert als Normalisierung des CRP auf < 10 mg/l ohne sukzessive Anstiege auf ≥ 10 mg/l.</p> <p>Ad 4. Erfolgreiches Prednison-Ausschleichen wurde definiert als ausbleibender Bedarf an Notfallmedikation. Der Gebrauch von zusätzlichem Prednison mit einer kumulativen Dosis von ≤ 100 mg (oder Äquivalent), d. h., über die in der Ausschleichphase vorgesehene Dosis hinaus, beispielsweise zur Behandlung von nicht im Zusammenhang mit der PMR stehenden UE, war möglich.</p> <p>Die Ergebnisse werden in Form von Responderanalysen dargestellt. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden für die anhaltende Remission zu Woche 52 auf Basis aller vier Einzelkomponenten und für die Einzelkomponenten mittels logistischer Regression berechnet.</p> <p>Die Analyse basierte auf den Daten der ITT-Population, die alle randomisierten Patienten einschloss. Für die Primäranalyse der anhaltenden Remission zu Woche 52 wurden Patienten, die die Studie vor Woche 52 beendeten und vor dem Abbruch keinen Schub erlitten, als Non-Responder imputiert. Fehlende Werte der Einzelkomponenten wurden nicht imputiert.</p>
<p>CRP: C-reaktives Protein; CS: Corticosteroids; ESR: Erythrozyten-Sedimentationsrate; ITT: Intention-To-Treat; KI: Konfidenzintervall; OR: Odds Ratio; PMR: Polymyalgia rheumatica; RD: Risikodifferenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis.</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-24: Bewertung des Verzerrungspotenzials für den Endpunkt Anhaltende Remission zu Woche 52 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Anhaltende Remission zu Woche 52 ist als niedrig einzustufen, da eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Anhaltende Remission zu Woche 52 als niedrig eingestuft (Tabelle 4-24).

Stellen Sie die Ergebnisse für den Endpunkt Anhaltende Remission zu Woche 52 für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-25: Ergebnisse für Anhaltende Remission zu Woche 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
17/60 (28,3)	6/58 (10,3)	3,43 [1,23; 9,55] 0,0190	2,74 [1,15; 6,52] 0,0227	17,99 [4,01; 31,97] 0,0121
Patienten, die die Studie vor Woche 52 beendeten und vor dem Abbruch keinen Schub erlitten, wurden als Non-Responder imputiert. CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko. Quelle: Sanofi 2024e				

Zum Studienende nach 52 Behandlungswochen hielt die Remission bei 17 Patienten (28,3 %) im Sarilumab-Arm an, verglichen mit 6 Patienten (10,3 %) im Kontrollarm. Die Unterschiede zwischen den Studienarmen fielen statistisch signifikant und klinisch relevant zugunsten von Sarilumab + CST aus (Tabelle 4-25).

Der für den Endpunkt anhaltende Remission zu Woche 52 beobachtete statistisch signifikante und klinisch relevante Vorteil von Sarilumab + CST gegenüber Placebo + CST zeigte sich, mit Ausnahme von Remission spätestens zu Woche 12, ebenfalls für die Einzelkomponenten (Tabelle 4-26).

Tabelle 4-26: Ergebnisse der Einzelkomponenten für Anhaltende Remission zu Woche 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Remission spätestens zu Woche 12^a				
28/60 (46,7)	22/58 (37,9)	1,43 [0,68; 3,01] 0,3397	1,23 [0,80; 1,89] 0,3454	8,74 [-9,21; 26,68] 0,3369
Kein Krankheitsschub von Woche 12 bis 52^b				
33/60 (55,0)	19/58 (32,8)	2,51 [1,18; 5,34] 0,0175	1,68 [1,08; 2,60] 0,0206	22,24 [4,61; 39,87] 0,0139
Anhaltende CRP-Reduktion von Woche 12 bis 52^c				
40/60 (66,7)	26/58 (44,8)	2,46 [1,16; 5,23] 0,0195	1,49 [1,06; 2,09] 0,0223	21,84 [4,16; 39,52] 0,0159
Erfolgreiches Prednison-Ausschleichen von Woche 12 bis 52^d				
30/60 (50,0)	14/58 (24,1)	3,14 [1,42; 6,95] 0,0051	2,07 [1,22; 3,51] 0,0068	25,86 [8,91; 42,81] 0,0031
Fehlende Werte wurden nicht imputiert. a: definiert als Auflösung der Zeichen und Symptome der PMR und CRP-Normalisierung (< 10 mg/l). b: definiert als Fehlen eines Wiederauftretens von Zeichen und Symptomen oder eines Anstiegs der ESR, im Zusammenhang mit aktiver PMR, die eine Erhöhung der CS-Dosis erforderlich machte. c: definiert als CRP < 10 mg/l ohne nachfolgende Anstiege auf ≥ 10 mg/l bei zwei oder mehr Visiten nach Woche 12 bis Woche 52. d: definiert als kein Bedarf an Notfallmedikation. CRP: C-reaktives Protein; CS: Corticosteroids; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ESR: Erythrozyten-Sedimentationsrate; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; PMR: Polymyalgia rheumatica; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko. Quelle: Sanofi 2024e				

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt Anhaltende Remission zu Woche 52 auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.3 Endpunkt Kumulative Corticosteroid-Gesamtdosis über 52 Wochen - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-27: Operationalisierung des Endpunkts Kumulative CS-Gesamtdosis

Studie	Operationalisierung
SAPHYR	<p>Die kumulative CS-Gesamtdosis wurde als Differenz zwischen tatsächlicher und erwarteter Gesamtdosis nach 52 Behandlungswochen zusammengefasst. Der erwarteten kumulativen Dosis liegt das protokollgerechte Prednison-Ausschleichen im Behandlungsverlauf zugrunde, während die tatsächliche kumulative Dosis neben dem protokollgerechten Prednison-Ausschleichen auch die Gabe von add-on-Prednison bis Woche 12 und Notfallmedikation von Woche 12 bis 52 berücksichtigt.</p> <p>Es wurden Median, einschließlich Min und Max, mit 95 %-KI und zugehörigem p-Wert anhand des nichtparametrischen Wilcoxon-Rangsummentests bestimmt.</p> <p>Die Analyse basierte auf den Daten der ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten.</p>
CS: Corticosteroids; ITT: Intention-To-Treat; KI: Konfidenzintervall; Max: Maximum; Min: Minimum.	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-28: Bewertung des Verzerrungspotenzials für den Endpunkt Kumulative CS-Gesamtdosis in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Kumulative CS-Gesamtdosis ist als niedrig einzustufen, da eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss, die mindestens eine Dosis der Studienmedikation erhalten hatten. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Kumulative CS-Gesamtdosis als niedrig eingestuft (Tabelle 4-28).

Stellen Sie die Ergebnisse für den Endpunkt Kumulative CS-Gesamtdosis für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-29: Ergebnisse für Kumulative CS-Gesamtdosis aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Endpunkt	Sarilumab + CST (N=59) ^a	Placebo + CST (N=58)	p-Wert ^b
Erwartete Gesamtdosis Median (mg) [95 %-KI] ^c (Min : Max)	777,0 [777,0; 777,0] (224 : 777)	2.044,0 [2.031,0; 2.044,0] (404 : 2.044)	NA
Tatsächliche Gesamtdosis Median (mg) [95 %-KI] ^d (Min : Max)	777,0 [777,0; 777,0] (224 : 2.885)	2.044,0 [2.044,0; 2.455,0] (404 : 3.676)	< 0,0001
Differenz Median (mg) [95 %-KI] (Min : Max)	0,0 [0,0; 0,0] (-13 : 2.108)	199,5 [0,0; 496,0] (-57 : 1.632)	0,0189

a: Ein Patient erhielt keine Studienmedikation.
b: p-Wert des Wilcoxon-Rangsummentests.
c: Protokollgerechtes Prednison-Ausschleichen im Behandlungsverlauf.
d: Erwartete Prednison-Gesamtdosis inklusive add-on-Prednison und Notfallmedikation.
CS: Corticosteroids; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat;
KI: Konfidenzintervall; NA: Nicht anwendbar; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial).
Quelle: Sanofi 2021a

Sowohl die erwartete als auch die tatsächliche kumulative CS-Gesamtdosis wiesen im Median das gleiche Ergebnis auf. Sie fielen mit 2.044 mg im Kontrollarm gegenüber 777 mg im Sarilumab-Arm deutlich höher aus. Der Median der Differenz zwischen tatsächlicher und erwarteter kumulativer CS-Gesamtdosis war im Sarilumab-Arm niedriger als im Kontrollarm. Der Unterschied war statistisch signifikant (Tabelle 4-29).

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt Kumulative CS-Gesamtdosis auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.4 Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-30: Operationalisierung des Endpunkts Zeit bis zum ersten PMR-Schub nach klinischer Remission

Studie	Operationalisierung
SAPHYR	<p>Die Zeit bis zum ersten PMR-Schub wurde definiert als Zeitspanne zwischen Randomisierung und Auftreten des ersten PMR-Schubs nach Erreichen einer klinischen Remission. Patienten, die keine Remission erreichten, wurden als Patienten mit Schub zur Randomisierung betrachtet. Patienten mit Remission, die keinen Schub entwickelten, wurden zum Behandlungsende bis Woche 52 zensiert.</p> <p>Der Behandlungseffekt (HR mit zugehörigem zweiseitigem 95 %-KI und p-Wert) wurde durch eine Cox-Regression geschätzt. Neben Anzahl der Patienten mit Ausprägung wurde die mediane Dauer bis zum ersten Schub nach klinischer Remission, einschließlich 95 %-KI nach der Kaplan-Meier-Methode ermittelt unter Berücksichtigung der Brookmeyer-Crowley-Methode (log-log-Transformation) und zusätzlich grafisch dargestellt.</p> <p>Die Analyse basierte auf den Daten der ITT-Population, die alle randomisierten Patienten einschloss.</p>
HR: Hazard Ratio; ITT: Intention-To-Treat; KI: Konfidenzintervall; PMR: Polymyalgia rheumatica.	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-31: Bewertung des Verzerrungspotenzials für den Zeit bis zum ersten PMR-Schub nach klinischer Remission in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission ist als niedrig einzustufen, da eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Zeit bis zum ersten PMR-Schub nach klinischer Remission als niedrig eingestuft (Tabelle 4-31).

Stellen Sie die Ergebnisse für den Endpunkt Zeit bis zum ersten PMR-Schub für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-32: Ergebnisse für Zeit bis zum ersten PMR-Schub nach klinischer Remission aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Endpunkt	Sarilumab + CST (N=60)	Placebo + CST (N=58)	Sarilumab + CST versus Placebo + CST
Keine klinische Remission, n (%)	19 (31,7)	28 (48,3)	
Ereignis, n (%)	10 (16,7)	17 (29,3)	
Zensiert, n (%)	31 (51,7)	13 (22,4)	
Median, Tage [95 %-KI] ^a	NE [93,0; NE]	99,0 [1,0; 154,0]	
HR [95 %-KI] ^b			0,56 [0,35; 0,90]
p-Wert ^b			0,0153
Patienten ohne klinische Remission im Studienverlauf wurden als Patienten mit PMR-Schub zur Randomisierung betrachtet. a: Kaplan-Meier-Schätzer. b: Cox-Proportional-Hazard-Modell. CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); HR: Hazard Ratio; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; NE: Nicht berechenbar (Not Evaluable); PMR: Polymyalgia rheumatica; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial). Quelle: Sanofi 2021c			

Während der 52-wöchigen Behandlungsphase trat bei 16,7 % der Patienten im Sarilumab-Arm ein PMR-Schub nach Erreichen klinischer Remission auf, verglichen mit 29,3 % der Patienten im Kontrollarm. Der Median für die Zeit bis zum ersten PMR-Schub nach Erreichen klinischer Remission konnte im Sarilumab-Arm nicht geschätzt werden, während er im Kontrollarm bei 99,0 Tagen lag. Der Unterschied in der Zeit bis zum ersten PMR-Schub zwischen den beiden Studienarmen fiel statistisch signifikant zugunsten von Sarilumab + CST aus (Tabelle 4-32).

Die zugehörige Kaplan-Meier-Kurve ist in Abbildung 4-2 dargestellt.

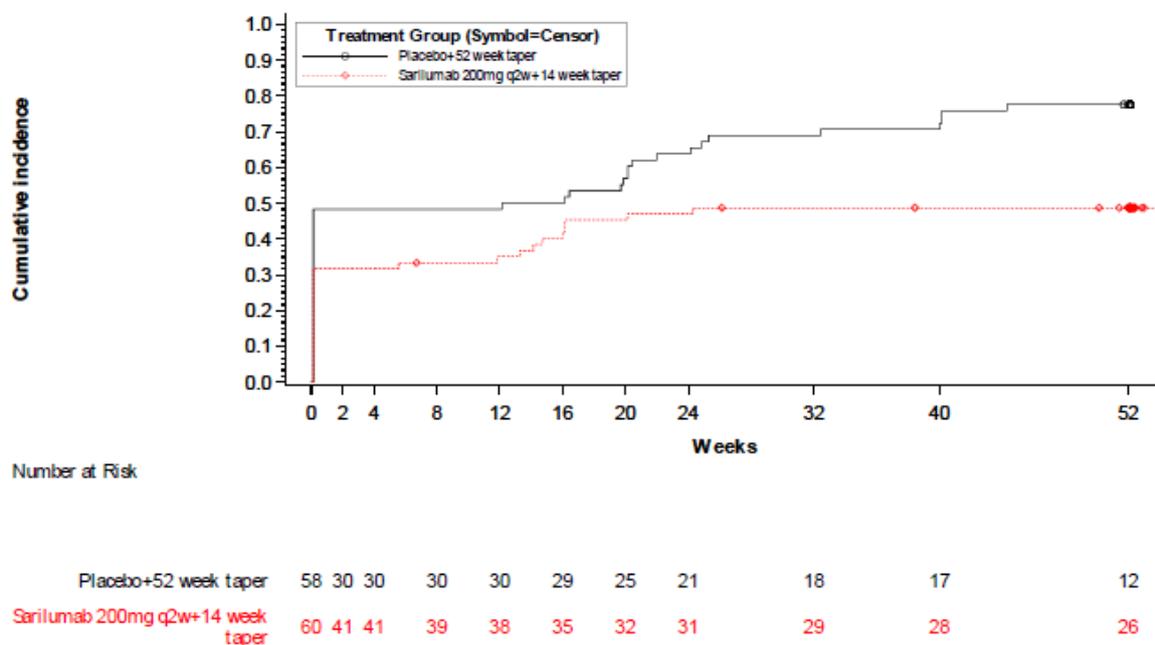


Abbildung 4-2: Kaplan-Meier-Kurve für den Endpunkt „Zeit bis zum ersten PMR-Schub nach klinischer Remission“ aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Die Zeit (Tage) wurde von der Randomisierung bis zum ersten PMR-Schub nach klinischer Remission bis Woche 52 berechnet.

ITT: Intention-To-Treat; PMR: Polymyalgia rheumatica; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial).

Quelle: Sanofi 2021c

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.5 Endpunkt PMR-Aktivitätsscore (PMR-AS) - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-33: Operationalisierung des Endpunkts PMR-AS

Studie	Operationalisierung
SAPHYR	<p>Der PMR-AS dient der Messung der Krankheitsaktivität. Er wurde aus den EULAR-Response-Kriterien für PMR entwickelt und umfasst folgende Einzelkomponenten:</p> <ul style="list-style-type: none"> ○ CRP (mg/dl) ○ VAS - Schmerz, Patientenbeurteilung (skaliert auf den Bereich 0 bis 10) ○ VAS - Krankheitsaktivität, Prüfarztbeurteilung (skaliert auf den Bereich 0 bis 10) ○ Dauer der Morgensteifigkeit (min x 0,1) ○ Beweglichkeit der oberen Gliedmaßen (EUL: 3 = keine, 2 = bis unterhalb Schultergürtel, 1 = bis Schultergürtel, 0 = bis oberhalb Schultergürtel) <p>Die Ergebnisse der einzelnen Parameter werden aufsummiert, wobei ein PMR-AS unter 7 eine niedrige, zwischen 7 und 17 eine mittlere und über 17 eine hohe Krankheitsaktivität anzeigt.</p> <p>Die Erhebung des PMR-AS erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Zur Bewertung der Schmerzsymptomatik durch den Patienten anhand einer VAS wurden die Ergebnisse der Fragebogenerhebung des HAQ-DI herangezogen, da sich beide VAS lediglich in ihrer Skalierung unterscheiden. Die Ergebnisse der Einzelkomponenten VAS – Schmerz, Patientenbeurteilung, und VAS – Krankheitsaktivität, Prüfarztbeurteilung, werden als Veränderung um $\geq -1,5$ Punkte gegenüber Baseline, entsprechend dem Responsekriterium ≥ 15 % der Skalenspannweite der VAS, dargestellt. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Die Ergebnisse für den globalen PMR-AS, die übrigen drei Einzelkomponenten sowie zusätzlich für die beiden VAS werden als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.</p> <p>Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung des PMR-AS und der Einzelkomponenten sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.</p>
<p>CRP: C-reaktives Protein; EUL: European League Against Rheumatism (EULAR); HAQ-DI: Health Assessment Questionnaire Disability Index; ITT: Intention-To-Treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Squares-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; OR: Odds Ratio; PMR-AS: Polymyalgia rheumatica-Aktivitätsscore; RD: Risikodifferenz; RR: Relatives Risiko; SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error); VAS: Visuelle Analogskala.</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-34: Bewertung des Verzerrungspotenzials für den Endpunkt PMR-AS in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	nein	ja	ja	hoch

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt PMR-AS ist als hoch einzustufen. Für den globalen PMR-AS und die Einzelkomponenten, für die die Veränderung zu Woche 52 gegenüber Baseline vorliegt, ist das Verzerrungspotenzial als hoch einzustufen, da die Analysen ausschließlich die randomisierten Patienten berücksichtigten, für die sowohl die Baseline- als auch eine Post-Baseline-Erhebung zu Woche 52 vorlagen. Das Verzerrungspotenzial für die Einzelkomponenten, für die Responderanalysen mittels VAS vorliegen, ist als niedrig einzustufen, da die Analysen auf Basis der ITT-Population erfolgten, die alle randomisierten Patienten einschloss. Eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen kann aufgrund des doppelblinden Studiendesigns ebenso ausgeschlossen werden wie eine ergebnisabhängige Berichterstattung, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Sonstige Aspekte für ein hohes Verzerrungspotenzial lagen nicht vor (Tabelle 4-34).

Stellen Sie die Ergebnisse für den Endpunkt PMR-AS für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-35: Veränderung des PMR-AS zu Woche 52 gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

Sarilumab + CST (N=60)			Placebo + CST (N=58)			Sarilumab + CST versus Placebo + CST	
n	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	n	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	LS-MWD [95 %-KI] ^a	p-Wert ^a
31	17,55 (8,65)	-15,57 (0,94)	28	20,82 (28,46)	-10,27 (0,97)	-5,30 [-8,01; -2,60]	0,0002
Hedges' g [95 %-KI] -1,01 [-1,52; -0,49]							
Fehlende Werte wurden nicht imputiert.							
a: MMRM mit Veränderung gegenüber Baseline als abhängige Variable und Baseline-Wert, Behandlungsgruppe, Visite und Interaktionen zwischen Baseline-Wert und Visite sowie zwischen Behandlung und Visite als Kovariaten.							
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Square-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit einer Erhebung zu Baseline und zu Woche 52; PMR-AS: Polymyalgia rheumatica-Aktivitätsscore; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error).							
Quellen: Sanofi 2024c							

Im Sarilumab-Arm verbesserte sich der globale PMR-AS zu Woche 52 gegenüber Baseline mit -15,57 Punkten deutlich stärker als im Kontrollarm mit -10,27 Punkten. Die Unterschiede zwischen den Studienarmen fielen mit einer LS-Mittelwertdifferenz von -5,30 statistisch signifikant und klinisch relevant zugunsten von Sarilumab + CST aus (Tabelle 4-35).

Das zugehörige Longitudinalplot ist in Abbildung 4-3 dargestellt.

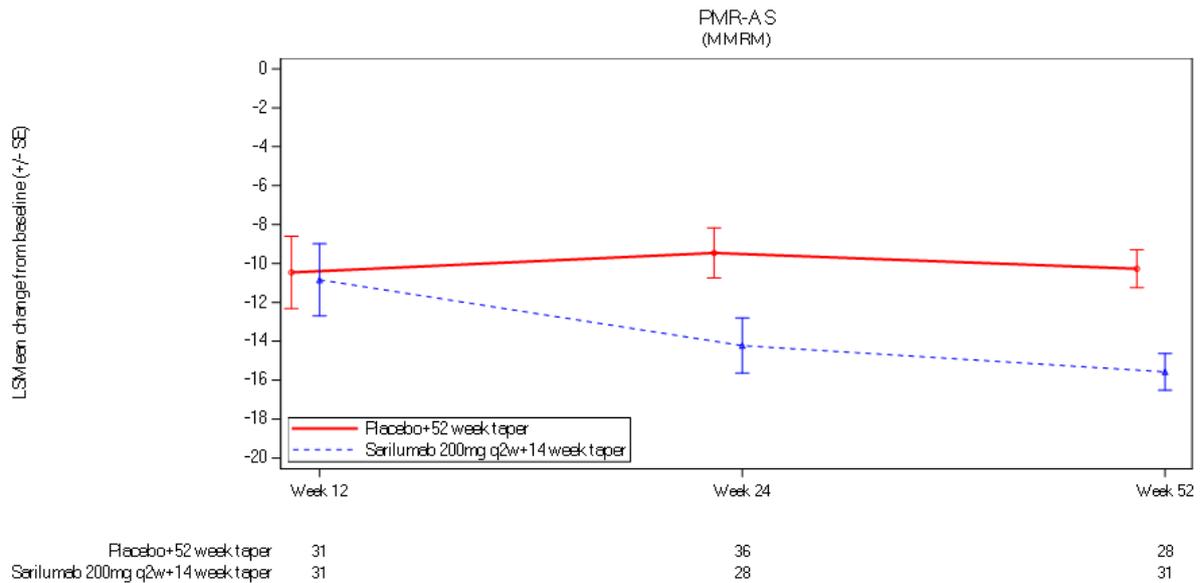


Abbildung 4-3: Longitudinalplot für die Veränderung des PMR-AS gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

LSMean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; PMR-AS: Polymyalgia rheumatica-Aktivitätsscore; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

Der für den globalen PMR-AS beobachtete statistisch signifikante Vorteil von Sarilumab + CST gegenüber Placebo + CST zu Woche 52 zeigte sich ebenfalls für die Einzelkomponenten CRP, Dauer der Morgensteifigkeit und Beweglichkeit der oberen Gliedmaßen. Dabei fielen die Unterschiede zwischen den Studienarmen ausschließlich für die CRP-Komponente auch klinisch relevant zugunsten von Sarilumab + CST aus (Tabelle 4-36).

Tabelle 4-36: Veränderung der Einzelkomponenten des PMR-AS zu Woche 52 gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

Sarilumab + CST (N=60)			Placebo + CST (N=58)			Sarilumab + CST versus Placebo + CST	
n	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	n	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	LS-MWD [95 %-KI] ^a Hedges' g [95 %-KI]	p-Wert ^a
CRP (mg/dl)							
43	10,35 (10,02)	-6,31 (1,67)	35	9,81 (12,27)	0,79 (1,76)	-7,10 [-11,91; -2,29] -0,66 [-1,10; -0,21]	0,0042
Dauer der Morgensteifigkeit (min x 0,1)							
48	66,35 (64,86)	-75,61 (5,87)	46	106,30 (216,84)	-53,18 (5,98)	-22,43 [-39,09; -5,77] -0,55 [-0,96; -0,14]	0,0088
Beweglichkeit der oberen Gliedmaßen (EUL [3-0])							
48	0,52 (0,80)	-0,47 (0,06)	46	0,46 (0,62)	-0,23 (0,06)	-0,24 [-0,40; -0,08] -0,60 [-1,00; -0,19]	0,0044
Fehlende Werte wurden nicht imputiert.							
a: MMRM mit Veränderung gegenüber Baseline als abhängige Variable und Baseline-Wert, Behandlungsgruppe, Visite und Interaktionen zwischen Baseline-Wert und Visite sowie zwischen Behandlung und Visite als Kovariaten.							
CRP: C-reaktives Protein; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); EUL: European League Against Rheumatism (EULAR); KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Square-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit einer Erhebung zu Baseline und zu Woche 52; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error).							
Quellen: Sanofi 2024c							

Die zugehörigen Longitudinalplots sind in Abbildung 4-4, Abbildung 4-5, Abbildung 4-6 dargestellt.

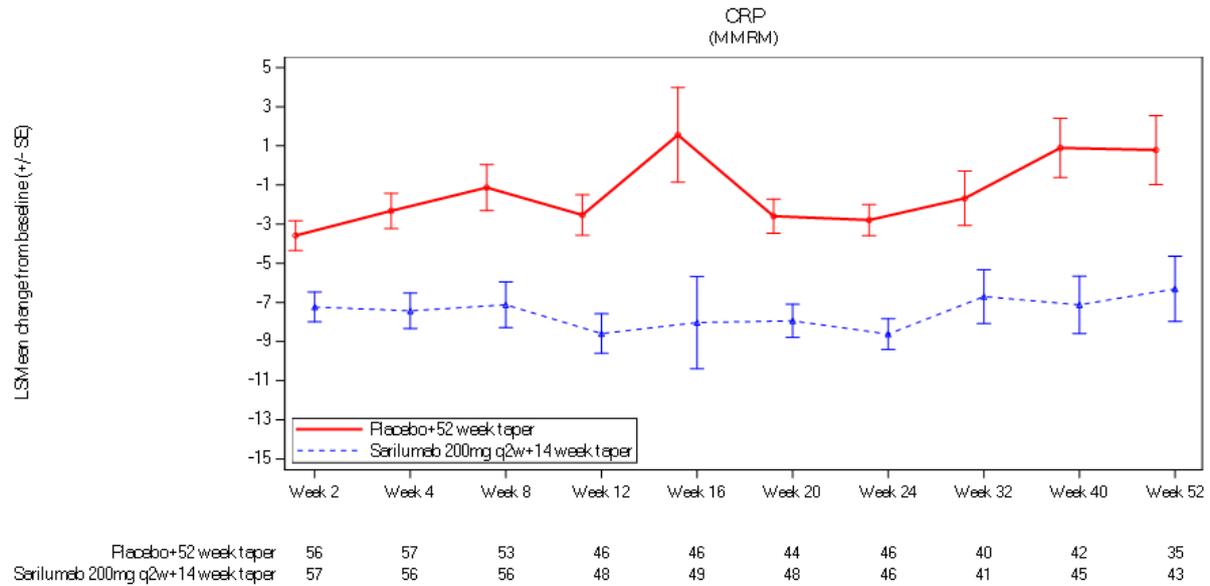


Abbildung 4-4: Longitudinalplot für die Veränderung des CRP gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

CRP: C-reaktives Protein; LSMean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

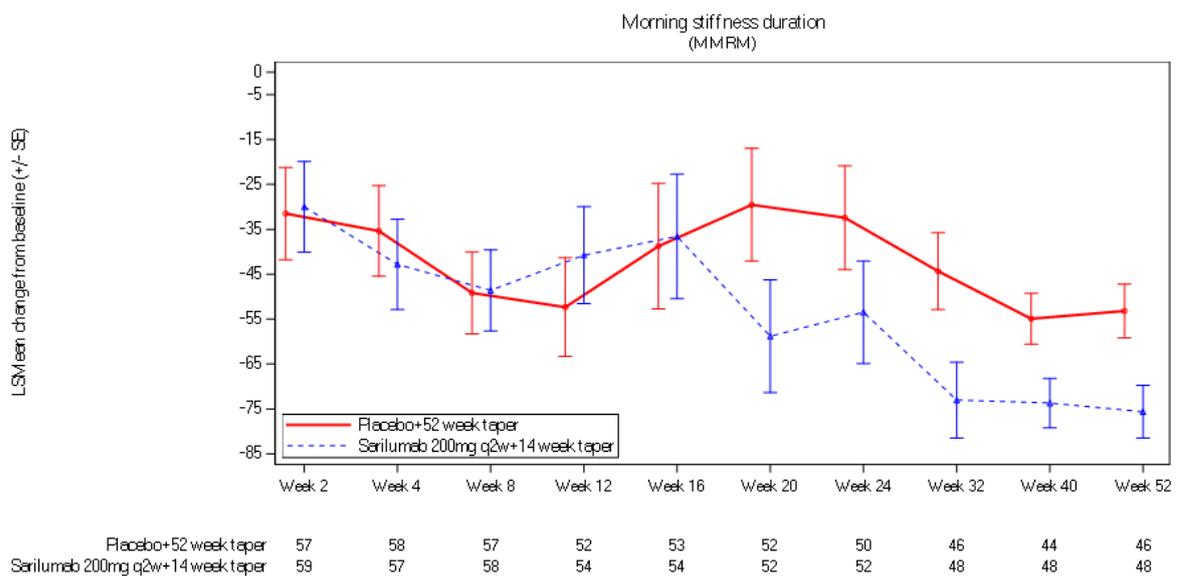


Abbildung 4-5: Longitudinalplot für die Veränderung der Dauer der Morgensteifigkeit gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

LSMean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

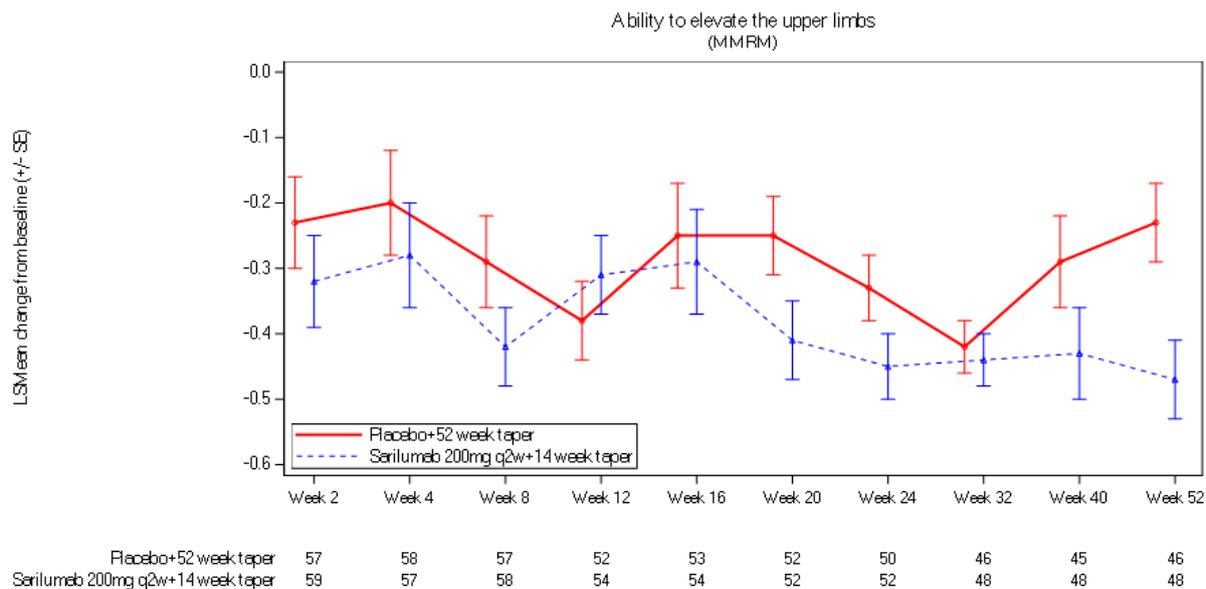


Abbildung 4-6: Longitudinalplot für die Veränderung der Beweglichkeit der oberen Gliedmaßen gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

LSMean: Kleinste-Quadrate-Mittelwert (Least Square Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

Tabelle 4-37: Rücklaufquoten der VAS – Schmerz, Patientenbeurteilung aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Sarilumab + CST N=60		Placebo + CST N=58	
	n1	n2 (%)	n1	n2 (%)
Baseline	59	58 (98,3)	58	58 (100,0)
Woche 12	59	46 (78,0)	58	47 (81,0)
Woche 24	52	42 (80,8)	54	43 (79,6)
Woche 52	48	43 (89,6)	48	35 (72,9)

CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat; N: Anzahl der Patienten; n1: Anzahl der Patienten mit erwarteter Erhebung; n2: Anzahl der Patienten mit vorhandener Erhebung; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); VAS: Visuelle Analogskala.
Quelle: Sanofi 2024e

Die Rücklaufquoten der VAS – Schmerz, Patientenbeurteilung lagen in beiden Studienarmen zu allen Erhebungszeitpunkten bei über 70 % (Tabelle 4-37).

Tabelle 4-38: Veränderung der Einzelkomponente VAS um $\geq -1,5$ Punkte gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
VAS - Schmerz, Patientenbeurteilung (0 bis 10)				
20/60 (33,3)	20/58 (34,5)	0,95 [0,44; 2,05] 0,8953	0,97 [0,58; 1,61] 0,8962	-1,15 [-18,41; 16,12] 0,8953
VAS - Krankheitsaktivität, Prüfarztbeurteilung (0 bis 10)				
30/60 (50,0)	21/58 (36,2)	1,76 [0,84; 3,71] 0,1346	1,38 [0,90; 2,12] 0,1408	13,79 [-4,09; 31,67] 0,1292
Patienten mit fehlenden Daten wurden als Non-Responder imputiert. CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; VAS: Visuelle Analogskala. Quelle: Sanofi 2024e				

Sowohl im Sarilumab-Arm als auch im Kontrollarm gab ein Drittel der Patienten eine Verbesserung des Schmerzempfindens um mindestens 1,5 Punkte an. Die Unterschiede zwischen den Studienarmen fielen somit nicht statistisch signifikant aus. Eine Verbesserung der Krankheitsaktivität um mindestens 1,5 Punkte gegenüber Baseline wurde von den Prüfarzten für die Hälfte der Patienten im Sarilumab-Arm berichtet, verglichen mit etwa einem Drittel im Kontrollarm. Es zeigte sich ein numerischer, aber nicht statistisch signifikanter, Vorteil für Sarilumab + CST (Tabelle 4-38).

Die Ergebnisse der Veränderung der VAS, Patienten- und Prüfarztbeurteilung, zu Woche 52 gegenüber Baseline mit den zugehörigen Longitudinalplots befinden sich in Anhang 4-G.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt PMR-AS auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.6 Patientenberichtete Endpunkte

4.3.1.3.1.6.1 Gesundheitszustand gemäß FACIT-F - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-39: Operationalisierung des Endpunkts Gesundheitszustand gemäß FACIT-F

Studie	Operationalisierung
SAPHYR	<p>Der FACIT-F wurde als befindlichkeitsspezifische Subskala des generischen Fragebogens FACIT-G zur Erfassung des Erschöpfungsgrads der Patienten angewandt. Er beinhaltet 13 Fragen, mit denen der Grad der Erschöpfung in den vergangenen 7 Tagen über eine fünfstufige Likert-Skala von 0 bis 4 (0 = überhaupt nicht, 1 = ein wenig, 2 = mäßig, 3 = ziemlich, 4 = sehr) gemessen wird. Die individuellen Antwortwerte werden auf einen Gesamtwert von 0 bis 52 summiert, wobei die Antwortwerte der 11 negativ formulierten Fragen umgekehrt werden. Ein höherer Gesamtwert bedeutet einen besseren Gesundheitszustand.</p> <p>Die Erhebung des Fragebogens erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Die Ergebnisse werden zum einen als Veränderung um $\geq 7,8$ Punkte, entsprechend dem Responsekriterium $\geq 15\%$ der Skalenspannweite des FACIT-F, berichtet. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Zum anderen werden die Ergebnisse als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.</p> <p>Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung des FACIT-F sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.</p>
<p>FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; FACIT-G: FACIT-General; ITT: Intention-To-Treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Square-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error).</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-40: Bewertung des Verzerrungspotenzials für den Endpunkt Gesundheitszustand gemäß FACIT-F in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Gesundheitszustand gemäß FACIT-F ist als niedrig einzustufen, da die Analyse auf Basis der ITT-Population erfolgte, die alle randomisierten Patienten einschloss. Eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen kann aufgrund des doppelblinden Studiendesigns ebenso ausgeschlossen werden wie eine ergebnisabhängige Berichterstattung, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Sonstige Aspekte für ein hohes Verzerrungspotenzial lagen nicht vor (Tabelle 4-40).

Stellen Sie die Ergebnisse für den Endpunkt Gesundheitszustand gemäß FACIT-F für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-41: Rücklaufquoten des FACIT-F aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Sarilumab + CST N=60		Placebo + CST N=58	
	n1	n2 (%)	n1	n2 (%)
Baseline	60	59 (98,3)	58	58 (100,0)
Woche 12	59	47 (79,7)	58	47 (81,0)
Woche 24	52	42 (80,8)	54	43 (79,6)
Woche 52	48	43 (89,6)	48	36 (75,0)

CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; ITT: Intention-To-Treat; N: Anzahl der Patienten; n1: Anzahl der Patienten mit erwarteter Erhebung; n2: Anzahl der Patienten mit vorhandener Erhebung; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial).

Quelle: Sanofi 2024e

Die Rücklaufquoten des FACIT-F lagen in beiden Studienarmen zu allen Erhebungszeitpunkten bei über 70 % (Tabelle 4-41).

Tabelle 4-42: Veränderung des FACIT-F um $\geq 7,8$ Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
		OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
n/N (%)	n/N (%)			
24/60 (40,0)	17/58 (29,3)	1,61 [0,74; 3,49] 0,2267	1,36 [0,82; 2,28] 0,2331	10,69 [-6,55; 27,92] 0,2218

Patienten mit fehlenden Daten wurden als Non-Responder imputiert.
 CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue Scale; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko.
 Quelle: Sanofi 2024e

Im Sarilumab-Arm erreichten 24 Patienten (40,0 %) eine Verbesserung des patientenempfindenen Erschöpfungszustands gemäß FACIT-F um mindestens 7,8 Punkte, verglichen mit 17 Patienten (29,3 %) im Kontrollarm. Daraus ergab sich ein numerischer, aber nicht statistisch signifikanter, Vorteil für Sarilumab + CST (Tabelle 4-42).

Die Ergebnisse der Veränderung des FACIT-F zu Woche 52 gegenüber Baseline mit dem zugehörigen Longitudinalplot befinden sich in Anhang 4-G.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt Gesundheitszustand gemäß FACIT-F auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.6.2 Gesundheitszustand gemäß EQ-5D VAS - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-43: Operationalisierung des Endpunkts Gesundheitszustand gemäß EQ-5D VAS

Studie	Operationalisierung
SAPHYR	<p>Der EQ-5D ist ein generischer Fragebogen zur Erfassung des Gesundheitszustands der Patienten.</p> <p>Anhand der VAS des EQ-5D ordnet der Patient seinem aktuellen Gesundheitszustand einen Wert zwischen 0 (denkbar schlechtester Gesundheitszustand) und 100 (denkbar bester Gesundheitszustand) zu.</p> <p>Die Erhebung erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Die Ergebnisse werden zum einen als Veränderung um ≥ 15 Punkte, entsprechend dem Responsekriterium $\geq 15\%$ der Skalenspannweite der EQ-5D VAS, berichtet. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Zum anderen werden die Ergebnisse als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.</p> <p>Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung der EQ-5D VAS sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.</p>
<p>EQ-5D: European Quality of Life 5 Dimensions; ITT: Intention-To-Treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Square-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; OR: Odds Ratio; RD: Risikodifferenz; RR: Relatives Risiko; SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error); VAS: Visuelle Analogskala.</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-44: Bewertung des Verzerrungspotenzials für den Endpunkt Gesundheitszustand gemäß EQ-5D VAS in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Gesundheitszustand gemäß EQ-5D VAS ist als niedrig einzustufen, da die Analyse auf Basis der ITT-Population erfolgte, die alle randomisierten Patienten einschloss. Eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen kann aufgrund des doppelblinden Studiendesigns ebenso ausgeschlossen werden wie eine ergebnisabhängige Berichterstattung, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Sonstige Aspekte für ein hohes Verzerrungspotenzial lagen nicht vor (Tabelle 4-44).

Stellen Sie die Ergebnisse für den Endpunkt Gesundheitszustand gemäß EQ-5D VAS für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-45: Rücklaufquoten der EQ-5D VAS aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Sarilumab + CST N=60		Placebo + CST N=58	
	n1	n2 (%)	n1	n2 (%)
Baseline	59	58 (98,3)	58	58 (100,0)
Woche 12	59	46 (78,0)	58	47 (81,0)
Woche 24	52	42 (80,8)	54	43 (79,6)
Woche 52	48	43 (89,6)	48	35 (72,9)

CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); EQ-5D: European Quality of Life 5 Dimensions; ITT: Intention-To-Treat; N: Anzahl der Patienten; n1: Anzahl der Patienten mit erwarteter Erhebung; n2: Anzahl der Patienten mit vorhandener Erhebung; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); VAS: Visuelle Analogskala.
Quelle: Sanofi 2024e

Die Rücklaufquoten der EQ-5D VAS lagen in beiden Studienarmen zu allen Erhebungszeitpunkten bei über 70 % (Tabelle 4-45).

Tabelle 4-46: Veränderung der EQ-5D VAS um ≥ 15 Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
		OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
/N (%)	n/N (%)			
16/60 (26,7)	9/58 (15,5)	1,98 [0,79; 4,98] 0,1451	1,72 [0,82; 3,60] 0,1517	11,15 [-3,57; 25,86] 0,1361

Patienten mit fehlenden Daten wurden als Non-Responder imputiert.
 CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); EQ-5D: European Quality of Life 5 Dimensions;
 ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis;
 OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial);
 RD: Risikodifferenz; RR: Relatives Risiko; VAS: Visuelle Analogskala.
 Quelle: Sanofi 2024e

Im Sarilumab-Arm berichteten 16 Patienten (26,7 %) und damit fast doppelt so viele wie im Kontrollarm (9 Patienten, 15,5 %) eine Verbesserung des Gesundheitszustands gemäß EQ-5D VAS um mindestens 15 Punkte. Es zeigte sich ein numerischer, aber nicht statistisch signifikanter, Vorteil für Sarilumab + CST (Tabelle 4-46).

Die Ergebnisse der Veränderung der EQ-5D VAS zu Woche 52 gegenüber Baseline mit dem zugehörigen Longitudinalplot befinden sich in Anhang 4-G.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt Gesundheitszustand gemäß EQ-5D VAS auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.6.3 Körperlicher Funktionsstatus gemäß HAQ-DI - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-47: Operationalisierung des Endpunkts Körperlicher Funktionsstatus gemäß HAQ-DI

Studie	Operationalisierung
SAPHYR	<p>Der HAQ-DI wurde als spezifischer Fragebogen für rheumatische Erkrankungen zur Erfassung der körperlichen Mobilität bzw. zur Einschätzung der funktionellen Einschränkungen der Patienten anhand der folgenden 8 Kategorien angewandt:</p> <ul style="list-style-type: none"> ○ Ankleiden und Körperpflege ○ Aufstehen ○ Essen ○ Gehen ○ Hygiene ○ Erreichen von Gegenständen ○ Greifen ○ Andere Tätigkeiten <p>Jede Kategorie beinhaltet Fragen zu 2 oder 3 spezifischen Aktivitäten, mit denen der Patient seine funktionellen Einschränkungen in den vergangenen 7 Tagen beurteilt. Darüber hinaus wird die Benutzung von Hilfsmitteln (7 Fragen) bzw. die Notwendigkeit fremder Hilfe (4 Fragen) angegeben. Die Beantwortung der Fragen erfolgt über eine vierstufige Likert-Skala von 0 bis 3 (0 = ohne jede Schwierigkeit, 1 = mit einigen Schwierigkeiten, 2 = mit großen Schwierigkeiten, 3 = nicht dazu in der Lage). Zur Berechnung des HAQ-DI werden die höchsten Punktwerte aus den acht Kategorien gemittelt. Je höher der Index ausfällt, desto schwerer ist die Einschränkung im Alltag.</p> <p>Neben dem Standard-Score beinhaltet der HAQ-DI die Beurteilung des Schmerzempfindens und der globalen Krankheitsaktivität durch den Patienten mittels VAS auf einer Skala von 0 (keine Schmerzen bzw. sehr gute Gesundheit) bis 100 (sehr starke Schmerzen bzw. sehr schlechte Gesundheit).</p> <p>Die Erhebung des Fragebogens erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Die Ergebnisse für den Standard-Score und die Beurteilung der globalen Krankheitsaktivität werden zum einen als Veränderung um $\geq -0,45$ Punkte (Standard-Score) bzw. -15 Punkte (globale Krankheitsaktivität), entsprechend dem Responsekriterium $\geq 15\%$ der Skalenspannweite des HAQ-DI, berichtet. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Zum anderen werden die Ergebnisse sowohl für den Standard-Score als auch für die globale Krankheitsaktivität und die 8 Kategorien als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.</p> <p>Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung des HAQ-DI sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.</p> <p>Die Ergebnisse des Schmerzempfindens durch den Patienten mittels VAS sind in Abschnitt 4.3.1.3.1.5 zum Endpunkt PMR-AS dargestellt. Dabei wurde die VAS von 0 (keine Schmerzen) bis 10 (sehr starke Schmerzen) skaliert.</p>

Studie	Operationalisierung
	HAQ-DI: Health Assessment Questionnaire Disability Index; ITT: Intention-To-Treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Square-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; OR: Odds Ratio; PMR-AS: Polymyalgia rheumatica-Aktivitätsscore; RD: Risikodifferenz; RR: Relatives Risiko; SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error); VAS: Visuelle Analogskala.

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-48: Bewertung des Verzerrungspotenzials für den Endpunkt Körperlicher Funktionsstatus gemäß HAQ-DI in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Körperlicher Funktionsstatus gemäß HAQ-DI ist als niedrig einzustufen, da die Analyse auf Basis der ITT-Population erfolgte, die alle randomisierten Patienten einschloss. Eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen kann aufgrund des doppelblinden Studiendesigns ebenso ausgeschlossen werden wie eine ergebnisabhängige Berichterstattung, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Sonstige Aspekte für ein hohes Verzerrungspotenzial lagen nicht vor (Tabelle 4-48).

Stellen Sie die Ergebnisse für den Endpunkt Körperlicher Funktionsstatus gemäß HAQ-DI für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-49: Rücklaufquoten des HAQ-DI aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Sarilumab + CST N=60		Placebo + CST N=58	
	n1	n2 (%)	n1	n2 (%)
Baseline	59	58 (98,3)	58	58 (100,0)
Woche 12	59	46 (78,0)	58	47 (81,0)
Woche 24	52	42 (80,8)	54	43 (79,6)
Woche 52	48	43 (89,6)	48	35 (72,9)

CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); HAQ-DI: Health Assessment Questionnaire Disability Index; ITT: Intention-To-Treat; N: Anzahl der Patienten; n1: Anzahl der Patienten mit erwarteter Erhebung; n2: Anzahl der Patienten mit vorhandener Erhebung; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial).
Quelle: Sanofi 2024e

Die Rücklaufquoten des HAQ-DI lagen in beiden Studienarmen zu allen Erhebungszeitpunkten bei über 70 % (Tabelle 4-49).

Tabelle 4-50: Veränderung des Standard-Scores des HAQ-DI um $\geq -0,45$ Punkte und der globalen Krankheitsaktivität um ≥ -15 Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Standard-Score				
19/60 (31,7)	10/58 (17,2)	2,22 [0,92; 5,37] 0,0749	1,84 [0,93; 3,63] 0,0808	14,43 [-1,00; 29,85] 0,0666
Globale Krankheitsaktivität				
21/60 (35,0)	14/58 (24,1)	1,69 [0,75; 3,81] 0,2011	1,45 [0,81; 2,58] 0,2076	10,86 [-5,65; 27,37] 0,1951

Patienten mit fehlenden Daten wurden als Non-Responder imputiert.
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); HAQ-DI: Health Assessment Questionnaire Disability Index; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko.
Quelle: Sanofi 2024e

Im Sarilumab-Arm erreichten fast doppelt so viele Patienten (n=19) eine Verbesserung des körperlichen Funktionsstatus gemäß Standard-Score des HAQ-DI um mindestens 0,45 Punkte wie im Kontrollarm (n=10). Ein vergleichbares Ergebnis ergab die Veränderung der patientenbeurteilten globalen Krankheitsaktivität. Im Sarilumab-Arm verbesserten sich 21 Patienten um mindestens 15 Punkte, verglichen mit 14 Patienten im Kontrollarm. Für beide Endpunkte zeigte sich ein numerischer, aber nicht statistisch signifikanter, Vorteil für Sarilumab + CST (Tabelle 4-50).

Die Ergebnisse der Veränderung des Standard-Scores und der globalen Krankheitsaktivität des HAQ-DI zu Woche 52 gegenüber Baseline mit den zugehörigen Longitudinalplots befinden sich in Anhang 4-G.

Tabelle 4-51: Veränderung der Kategorien des HAQ-DI zu Woche 52 gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

Sarilumab + CST (N=60)			Placebo + CST (N=58)			Sarilumab + CST versus Placebo + CST	
n	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	n	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	LS-MWD [95 %-KI] ^a Hedges' g [95 %-KI]	p-Wert ^a
Ankleiden und Körperpflege							
42	1,02 (0,90)	-0,43 (0,12)	35	0,86 (0,88)	-0,19 (0,13)	-0,24 [-0,59; 0,11] -0,31 [-0,76; 0,14]	0,1787
Aufstehen							
42	1,10 (0,79)	-0,41 (0,11)	35	0,94 (0,87)	-0,22 (0,12)	-0,18 [-0,52; 0,15] -0,25 [-0,70; 0,20]	0,2809
Essen							
42	0,86 (0,90)	-0,37 (0,11)	35	0,83 (1,01)	-0,08 (0,12)	-0,28 [-0,60; 0,03] -0,40 [-0,85; 0,05]	0,0799
Gehen							
42	0,83 [0,76]	-0,14 (0,10)	35	0,74 (0,74)	-0,17 (0,11)	0,03 [-0,27; 0,33] 0,04 [-0,41; 0,49]	0,8460
Hygiene							
42	1,24 (0,98)	-0,44 (0,12)	35	0,77 (1,03)	0,11 (0,14)	-0,56 [-0,92; -0,19] -0,69 [-1,14; -0,23]	0,0034

Sarilumab + CST (N=60)			Placebo + CST (N=58)			Sarilumab + CST versus Placebo + CST	
Erreichen von Gegenständen							
42	1,55 (0,94)	-0,63 (0,12)	35	1,26 (0,98)	-0,35 (0,13)	-0,28 [-0,64; 0,08]	0,1266 -0,35 [-0,80; 0,10]
Greifen							
42	1,12 (0,94)	-0,46 (0,12)	35	0,86 (0,94)	-0,28 (0,13)	-0,18 [-0,52; 0,17]	0,3076 -0,23 [-0,68; 0,22]
Andere Tätigkeiten							
42	1,38 (0,91)	-0,33 (0,13)	35	1,37 (1,06)	-0,16 (0,13)	-0,17 [-0,54; 0,19]	0,3477 -0,21 [-0,66; 0,24]
Fehlende Werte wurden nicht imputiert. a: MMRM mit Veränderung gegenüber Baseline als abhängige Variable und Baseline-Wert, Behandlungsgruppe, Visite und Interaktionen zwischen Baseline-Wert und Visite sowie zwischen Behandlung und Visite als Kovariaten. CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); HAQ-DI: Health Assessment Questionnaire Disability Index; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Square-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit einer Erhebung zu Baseline und zu Woche 52; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error). Quellen: Sanofi 2024c							

In der Kategorie Hygiene des HAQ-DI fiel die Veränderung zu Woche 52 gegenüber Baseline zwischen den Studienarmen statistisch signifikant und klinisch relevant zugunsten von Sarilumab + CST aus. Die Kategorie Gehen zeigte für den Sarilumab-Arm einen numerischen, aber nicht statistisch signifikanten, Nachteil gegenüber dem Kontrollarm. Für die übrigen 6 Kategorien ergaben sich zwischen den Studienarmen zu Woche 52 gegenüber Baseline numerische, aber statistisch nicht signifikante, Vorteile für Sarilumab + CST (Tabelle 4-51).

Die zugehörigen Longitudinalplots sind in Abbildung 4-7, Abbildung 4-8, Abbildung 4-9, Abbildung 4-10, Abbildung 4-11, Abbildung 4-12, Abbildung 4-13 und Abbildung 4-14 dargestellt.

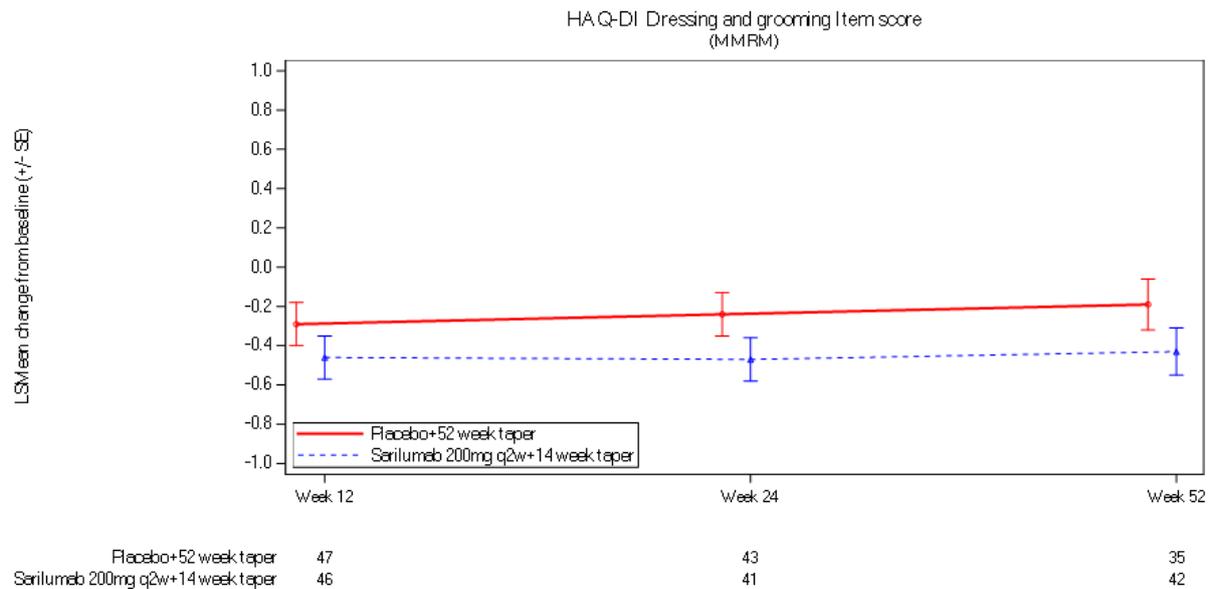


Abbildung 4-7: Longitudinalplot für die Veränderung der Kategorie Ankleiden und Körperpflege des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

HAQ-DI: Health Assessment Questionnaire Disability Index; LS Mean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

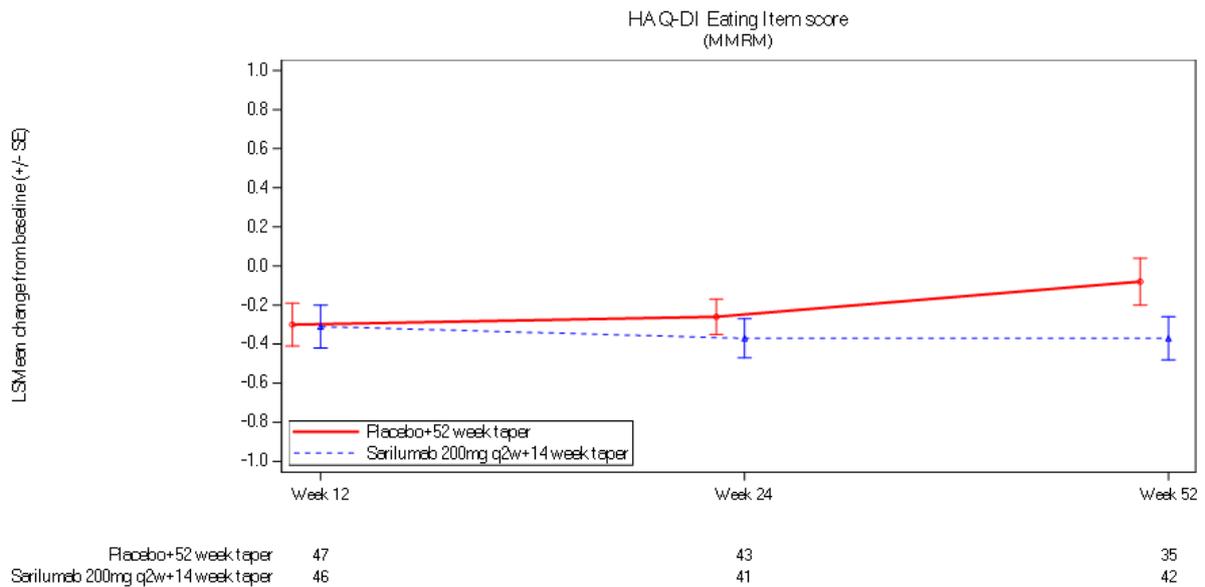


Abbildung 4-9: Longitudinalplot für die Veränderung der Kategorie Essen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

HAQ-DI: Health Assessment Questionnaire Disability Index; LSMean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

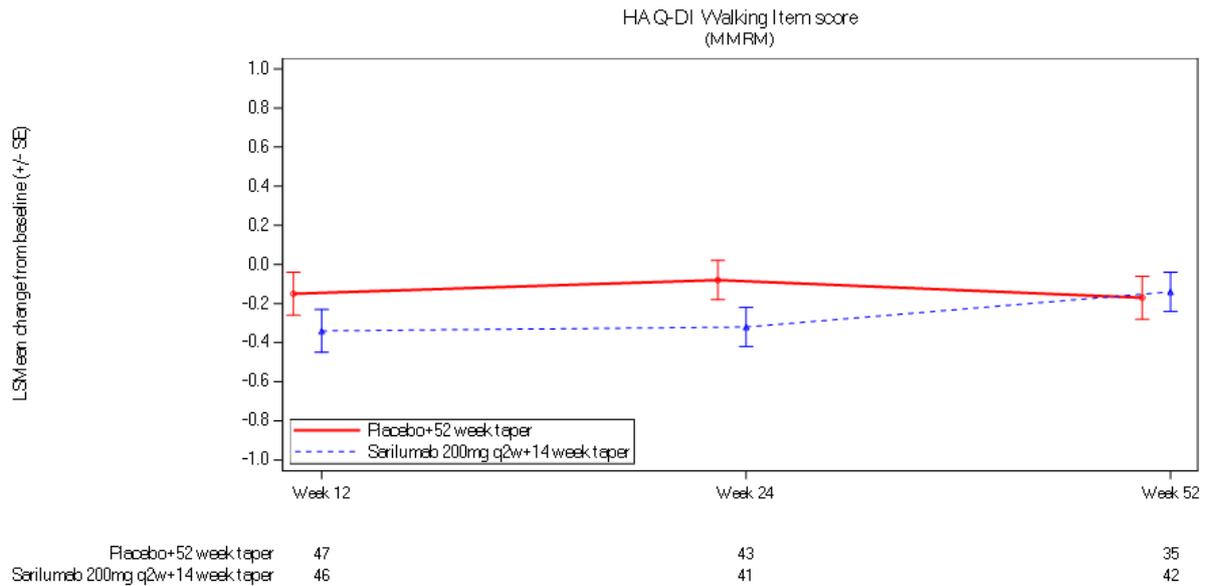


Abbildung 4-10: Longitudinalplot für die Veränderung der Kategorie Gehen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

HAQ-DI: Health Assessment Questionnaire Disability Index; LS Mean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

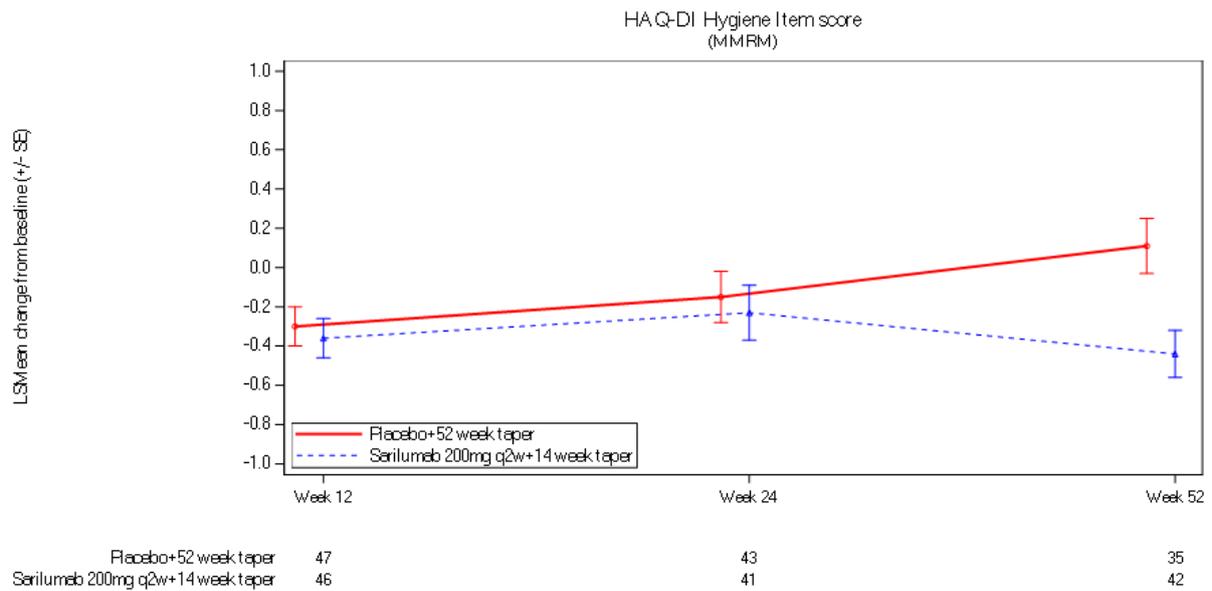


Abbildung 4-11: Longitudinalplot für die Veränderung der Kategorie Hygiene des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

HAQ-DI: Health Assessment Questionnaire Disability Index; LS Mean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

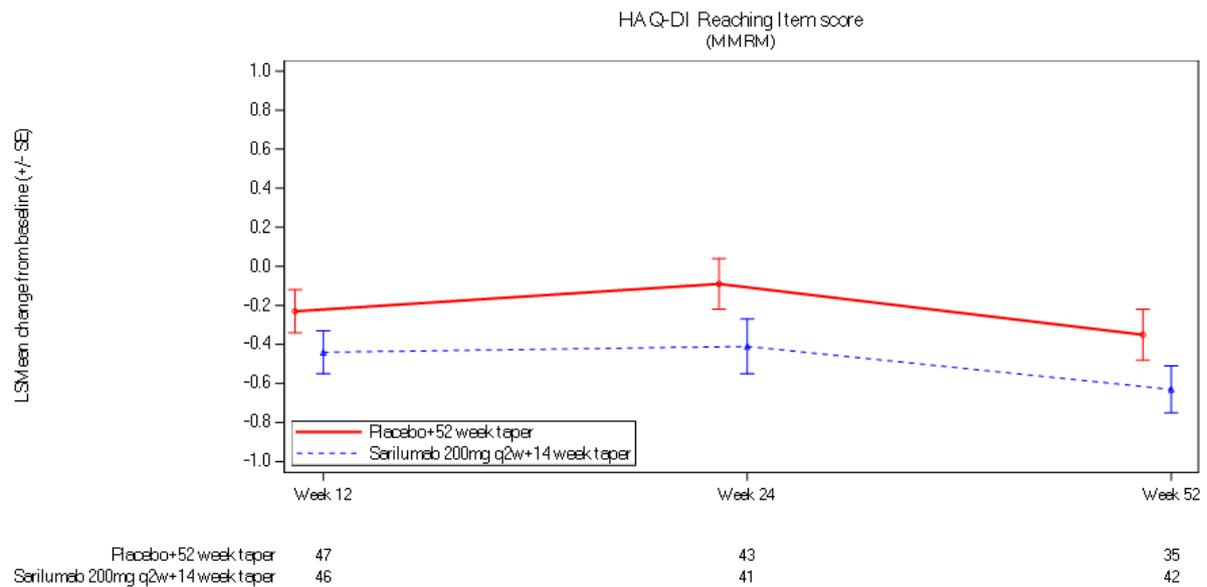


Abbildung 4-12: Longitudinalplot für die Veränderung der Kategorie Erreichen von Gegenständen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

HAQ-DI: Health Assessment Questionnaire Disability Index; LS Mean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

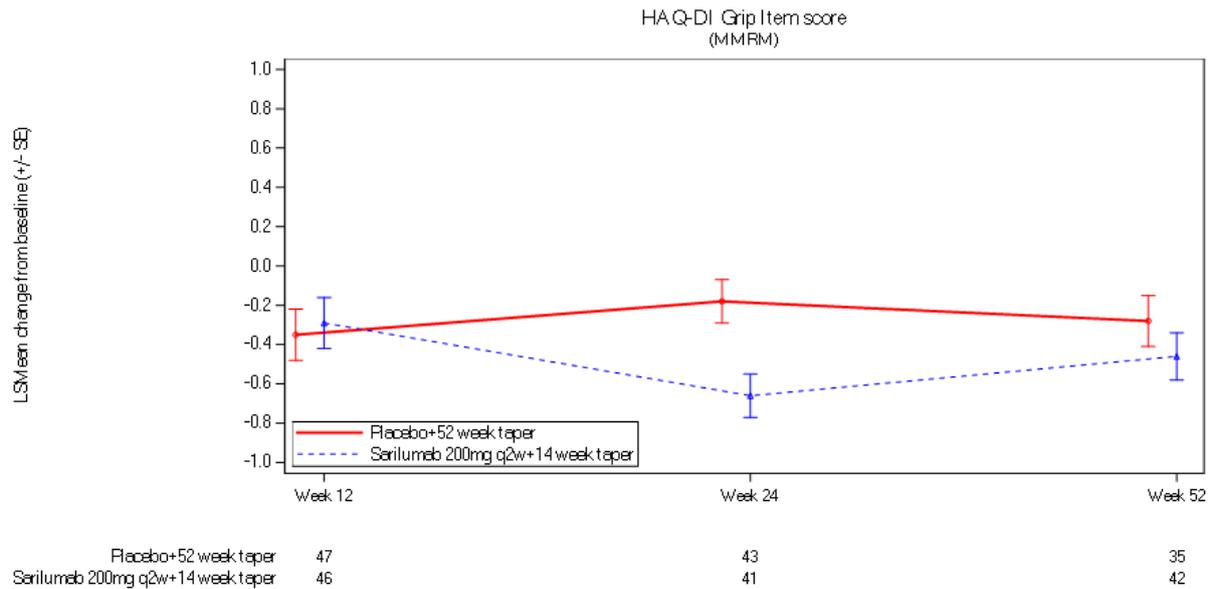


Abbildung 4-13: Longitudinalplot für die Veränderung der Kategorie Greifen des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

HAQ-DI: Health Assessment Questionnaire Disability Index; LS Mean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

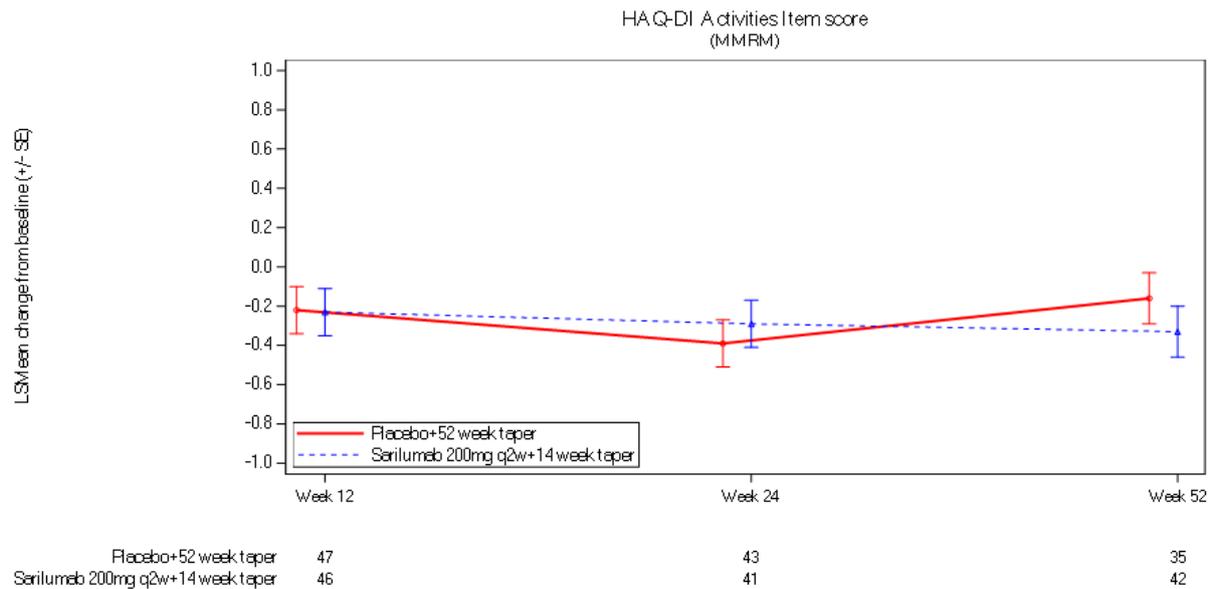


Abbildung 4-14: Longitudinalplot für die Veränderung der Kategorie Andere Tätigkeiten des HAQ-DI gegenüber Baseline aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

HAQ-DI: Health Assessment Questionnaire Disability Index; LS Mean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024c

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt Körperlicher Funktionsstatus gemäß HAQ-DI auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.6.4 Gesundheitsbezogene Lebensqualität gemäß SF-36v2 - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-52: Operationalisierung des Endpunkts Gesundheitsbezogene Lebensqualität gemäß SF-36v2

Studie	Operationalisierung
SAPHYR	<p>Der SF-36v2 wurde als generischer Fragebogen zur Erfassung der gesundheitsbezogenen Lebensqualität der Patienten anhand der beiden Hauptdimensionen PCS und MCS angewandt.</p> <p>Die PCS wird aus den folgenden Subskalen gebildet:</p> <ul style="list-style-type: none"> ○ Körperliche Funktionsfähigkeit ○ Körperliche Rollenfunktion ○ Körperliche Schmerzen ○ Allgemeine Gesundheitswahrnehmung <p>Die MCS wird aus folgenden Subskalen gebildet:</p> <ul style="list-style-type: none"> ○ Vitalität ○ Soziale Funktionsfähigkeit ○ Emotionale Rollenfunktion ○ Psychisches Wohlbefinden. <p>Pro Subskala ergeben 2 bis 10 Fragen insgesamt 36 Fragen, von denen eine Frage zur Veränderung der allgemeinen Gesundheitswahrnehmung in den vergangenen 12 Monaten in der Auswertung unberücksichtigt bleibt. Alle Fragen werden mit Hilfe einer drei- bis sechsstufigen Likert-Skala beantwortet und die erhaltenen Werte für jede Subskala in einen Bereich zwischen 0 (schlechteste Lebensqualität) und 100 (beste Lebensqualität) transformiert. PCS und MCS werden auf einen Mittelwert von 50 und eine Standardabweichung von 10 transformiert. Höhere Werte bedeuten eine bessere gesundheitsbezogene Lebensqualität.</p> <p>Die Erhebung des Fragebogens erfolgte zu Baseline und zu den Wochen 12, 24 und 52. Die Ergebnisse werden zum einen als Veränderung um ≥ 10 Punkte berichtet. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Zum anderen werden die Ergebnisse als Veränderung zu Woche 52 gegenüber Baseline anhand eines MMRM analysiert und zusätzlich mittels Longitudinalplot grafisch dargestellt. Es wurden MW (SD), LS-MW (SE) und LS-MWD mit 95 %-KI und p-Wert sowie Hedges' g mit 95 %-KI bestimmt.</p> <p>Die Responderanalysen basierten auf der ITT-Population, die alle randomisierten Patienten umfasste. Patienten mit fehlenden Daten wurden als Non-Responder imputiert. Für die Veränderung zu Woche 52 gegenüber Baseline wurden die randomisierten Patienten berücksichtigt, für die eine Erhebung der beiden Summenskalen des SF-36v2 sowohl zu Baseline als auch zu Woche 52 verfügbar war. Für die kontinuierliche Operationalisierung erfolgte keine Imputation fehlender Werte.</p>
<p>ITT: Intention-to-treat; KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Square-Mean Value Difference); MCS: Summenskala der psychischen Gesundheit (Mental Health Component Summary Score); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; OR: Odds Ratio; PCS: Summenskala der körperlichen Gesundheit (Physical Health Component Summary Score); RD: Risikodifferenz; RR: Relatives Risiko; SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error); SF-36v2: 36-Item Short Form Health Survey Version 2; SMD: Standardisierte Mittelwertdifferenz.</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-53: Bewertung des Verzerrungspotenzials für den Endpunkt Gesundheitsbezogene Lebensqualität gemäß SF-36v2 in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Gesundheitsbezogene Lebensqualität gemäß SF-36v2 ist als niedrig einzustufen, da die Analysen auf Basis der ITT-Population erfolgten, die alle randomisierten Patienten einschloss. Eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen kann aufgrund des doppelblinden Studiendesigns ebenso ausgeschlossen werden wie eine ergebnisabhängige Berichterstattung, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Sonstige Aspekte für ein hohes Verzerrungspotenzial lagen nicht vor (Tabelle 4-53).

Stellen Sie die Ergebnisse für den Endpunkt Gesundheitsbezogene Lebensqualität gemäß SF-36v2 für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-54: Rücklaufquoten des SF-36v2 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Zeitpunkt	Sarilumab + CST N=60		Placebo + CST N=58	
	n1	n2 (%)	n1	n2 (%)
Baseline	59	58 (98,3)	58	58 (100,0)
Woche 12	59	46 (78,0)	58	47 (81,0)
Woche 24	52	42 (80,8)	54	43 (79,6)
Woche 52	48	43 (89,6)	48	35 (72,9)

Zeitpunkt	Sarilumab + CST N=60		Placebo + CST N=58	
	n1	n2 (%)	n1	n2 (%)
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat; N: Anzahl der Patienten; n1: Anzahl der Patienten mit erwarteter Erhebung; n2: Anzahl der Patienten mit vorhandener Erhebung; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SF-36v2: 36-Item Short Form Health Survey Version 2. Quelle: Sanofi 2024e				

Die Rücklaufquoten des SF-36v2 lagen in beiden Studienarmen zu allen Erhebungszeitpunkten bei über 70 % (Tabelle 4-54).

Tabelle 4-55: Veränderung der PCS und MCS des SF-36v2 um ≥ 10 Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
PCS				
14/60 (23,3)	10/58 (17,2)	1,46 [0,58; 3,65] 0,4143	1,35 [0,65; 2,82] 0,4194	6,09 [-8,52; 20,70] 0,4106
MCS				
11/60 (18,3)	5/58 (8,6)	2,38 [0,76; 7,43] 0,1341	2,13 [0,78; 5,80] 0,1408	9,71 [-2,58; 22,01] 0,1204
Patienten mit fehlenden Daten wurden als Non-Responder imputiert. CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat; KI: Konfidenzintervall; MCS: Summenskala der psychischen Gesundheit (Mental Health Component Summary Score); N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; PCS: Summenskala der körperlichen Gesundheit (Physical Health Component Summary Score); RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; SF-36v2: 36-Item Short Form Health Survey Version 2. Quelle: Sanofi 2024e				

Im Sarilumab-Arm zeigte sich sowohl für die PCS als auch die MCS bei mehr Patienten (n = 14 bzw. n = 11) eine Verbesserung der gesundheitsbezogenen Lebensqualität anhand des SF-36v2 um mindestens 10 Punkte als im Kontrollarm (n = 10 bzw. n = 5). Daraus ergab sich für beide Subskalen ein numerischer, aber nicht statistisch signifikanter, Vorteil für Sarilumab + CST (Tabelle 4-55).

Die Ergebnisse der Veränderung der PCS und MCS des SF-36v2 zu Woche 52 gegenüber Baseline mit den zugehörigen Longitudinalplots befinden sich in Anhang 4-G.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Einzelergebnisse mithilfe von Meta-Analysen quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analysen (in der Regel als Forest-Plot) dar. Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie, warum eine Meta-Analyse durchgeführt wurde bzw. warum eine Meta-Analyse nicht durchgeführt wurde bzw. warum einzelne Studien ggf. nicht in die Meta-Analyse einbezogen wurden. Machen Sie auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Zur Übertragbarkeit der dargestellten Studienergebnisse für den Endpunkt Gesundheitsbezogene Lebensqualität gemäß SF-36v2 auf den deutschen Versorgungskontext wird auf Abschnitt 4.3.1.2.1 verwiesen.

4.3.1.3.1.7 Endpunkt Nebenwirkungen - RCT

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-56: Operationalisierung des Endpunkts Nebenwirkungen

Studie	Operationalisierung
SAPHYR	<p>Es wurden alle UE vom Zeitpunkt der ersten bis 60 Tage nach der letzten Verabreichung der Studienmedikation erfasst. Die Kodierung der UE erfolgte anhand der MedDRA-Version 24.0 hierarchisch von der entsprechenden SOC zum PT. SUE oder UE von besonderem Interesse, die zum Studienende weiter bestanden, wurden bis zur Auflösung oder Stabilisierung weiter beobachtet und dokumentiert. Zur Bewertung der Patientensicherheit während der Studie stand ein unabhängiger Sicherheitsbeauftragter zur Verfügung.</p> <p>Folgende sicherheitsrelevanten Endpunkte wurden berücksichtigt:</p> <ul style="list-style-type: none"> • UE gesamt • Schwere UE • SUE • UE, die zum Therapieabbruch führten • UE von besonderem Interesse <ul style="list-style-type: none"> ○ Schwangerschaft einer Studienpatientin oder Partnerin eines Studienpatienten ○ Symptomatische Überdosierung der Studienmedikation ○ ALT-Anstieg auf $\geq 3 \times$ ULN ○ Klinisch relevante Infektionen ○ Abnormale ALT- oder ANC-Werte oder Thrombozytopenie, die zum Behandlungsabbruch führten <p>Die Ergebnisse der sicherheitsrelevanten Endpunkte werden in Form von Responderanalysen dargestellt. OR, RR und RD mit den zugehörigen 95 %-KI und p-Werten wurden mittels logistischer Regression berechnet. Darüber hinaus werden UE gesamt, die bei mindestens 10 % der Patienten in einem Studienarm auftraten, sowie alle schweren UE und SUE, die bei mindestens 5 % der Patienten in einem Studienarm auftraten, in Form von Responderanalysen nach SOC und PT berichtet. UE, die zum Therapieabbruch führten, und UE von besonderem Interesse werden zusätzlich deskriptiv nach SOC und PT berichtet.</p> <p>Die Analyse basierte auf den Daten der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten.</p>
<p>ALT: Alanin-Aminotransferase; ANC: Absolute Neutrophilenzahl (Absolute Neutrophil Count); KI: Konfidenzintervall; MedDRA: Medizinisches Wörterbuch für Aktivitäten im Rahmen der Arzneimittelzulassung (Medical Dictionary for Regulatory Activities); OR: Odds Ratio; PT: Bevorzugter Begriff (Preferred Term); RD: Risikodifferenz; RR: Relatives Risiko; SOC: Systemorganklasse (System Organ Class); SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; ULN: Obere Normgrenze (Upper Limit of Normal).</p>	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-57: Bewertung des Verzerrungspotenzials für den Endpunkt Nebenwirkungen in RCT mit dem zu bewertenden Arzneimittel

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
SAPHYR	niedrig	ja	ja	ja	ja	niedrig

Begründen Sie für jede Studie die abschließende Einschätzung.

Das Verzerrungspotenzial für den Endpunkt Nebenwirkungen ist als niedrig einzustufen, da eine Einflussnahme des Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Nebenwirkungen als niedrig eingestuft (Tabelle 4-57).

Stellen Sie die Ergebnisse für den Endpunkt Nebenwirkungen für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-58: Ergebnisse der sicherheitsrelevanten Endpunkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
		OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
UE gesamt				
56/59 (94,9)	49/58 (84,5)	3,43 [0,87; 13,58] 0,0788	1,12 [0,99; 1,27] 0,0711	10,43 [-0,56; 21,42] 0,0626
Schwere UE				
5/59 (8,5)	7/58 (12,1)	0,67 [0,20; 2,29] 0,5249	0,70 [0,23; 2,11] 0,5289	-3,59 [-14,70; 7,51] 0,5228

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
		OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
SUE				
8/59 (13,6)	12/58 (20,7)	0,60 [0,22; 1,62] 0,3108	0,66 [0,29; 1,50] 0,3164	-7,13 [-20,88; 6,62] 0,3063
UE, die zum Therapieabbruch führten				
7/59 (11,9)	4/58 (6,9)	1,82 [0,50; 6,67] 0,3646	1,72 [0,53; 5,63] 0,3701	4,97 [-5,66; 15,60] 0,3565
UE von besonderem Interesse, gesamt				
7/59 (11,9)	5/58 (8,6)	1,43 [0,42; 4,85] 0,5657	1,38 [0,46; 4,14] 0,5695	3,24 [-7,84; 14,33] 0,5632
UE von besonderem Interesse, leicht				
1/59 (1,7)	1/58 (1,7)	0,98 [0,06; 16,58] 0,9903	0,98 [0,06; 15,80] 0,9904	-0,03 [-4,78; 4,72] 0,9903
UE von besonderem Interesse, moderat				
4/59 (6,8)	2/58 (3,4)	2,04 [0,35; 11,79] 0,4241	1,97 [0,37; 10,51] 0,4291	3,33 [-4,70; 11,37] 0,4132
UE von besonderem Interesse, schwer				
2/59 (3,4)	2/58 (3,4)	0,98 [0,13; 7,37] 0,9862	0,98 [0,14; 6,89] 0,9863	-0,06 [-6,71; 6,60] 0,9862
UE von besonderem Interesse, schwerwiegend				
3/59 (5,1)	2/58 (3,4)	1,50 [0,24; 9,51] 0,6644	1,47 [0,25; 8,66] 0,6673	1,64 [-5,75; 9,03] 0,6618
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis. Quelle: Sanofi 2024h				

Im Sarilumab-Arm wurden für UE gesamt bei 56 Patienten (94,9 %) sowie für UE, die zum Therapieabbruch führten, und für UE von besonderem Interesse bei jeweils 7 Patienten (11,9 %) geringfügig höhere Inzidenzen registriert als im Kontrollarm bei 49 (84,5 %), 4 (6,9 %) bzw. 5 Patienten (8,6 %). Die UE von besonderem Interesse nach Schweregrad fielen für leichte UE bei jeweils einem Patienten (1,7 %) und schwere UE bei jeweils 2 Patienten (3,4 %) in beiden Studienarmen gleich aus, für moderate UE bei 4 Patienten (6,8 %) gegenüber

2 Patienten (3,4 %) und schwerwiegende UE bei 3 Patienten (5,1 %) gegenüber 2 Patienten (3,4 %) im Sarilumab-Arm geringfügig höher aus als im Kontrollarm. Die Anzahl an Patienten mit schweren UE und SUE lag dagegen im Kontrollarm mit 7 (12,1 %) bzw. 12 (20,7 %) etwas höher als im Sarilumab-Arm mit 5 (8,5 %) bzw. 8 (13,6 %). Die Unterschiede zwischen den Studienarmen sind nicht statistisch signifikant (Tabelle 4-58).

Tabelle 4-59: Ergebnisse der UE gesamt nach SOC aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
		OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Infektionen und parasitäre Erkrankungen				
22/59 (37,3)	29/58 (50,0)	0,59 [0,28; 1,25] 0,1696	0,75 [0,49; 1,14] 0,1729	-12,71 [-30,73; 5,31] 0,1649
Erkrankungen des Blutes und des Lymphsystems				
15/59 (25,4)	5/58 (8,6)	3,61 [1,20; 18,85] 0,0224	2,95 [1,14; 7,66] 0,0268	16,80 [3,41; 30,20] 0,0144
Stoffwechsel- und Ernährungsstörungen				
4/59 (6,8)	6/58 (10,3)	0,63 [0,17; 2,40] 0,4948	0,66 [0,19; 2,23] 0,4958	-3,57 [-13,80; 6,67] 0,4916
Psychiatrische Erkrankungen				
11/59 (18,6)	13/58 (22,4)	0,79 [0,32; 1,97] 0,6150	0,83 [0,40; 1,72] 0,6154	-3,77 [-18,55; 11,01] 0,6144
Erkrankungen des Nervensystems				
11/59 (18,6)	15/58 (25,9)	0,66 [0,27; 1,60] 0,3514	0,72 [0,36; 1,45] 0,3535	-7,22 [-22,40; 7,97] 0,3484
Augenerkrankungen				
9/59 (15,3)	10/58 (17,2)	0,86 [0,32; 2,34] 0,7714	0,88 [0,39; 2,04] 0,7715	-1,99 [-15,50; 11,52] 0,7713
Gefäßerkrankungen				
9/59 (15,3)	10/58 (17,2)	0,86 [0,32; 2,34] 0,7714	0,88 [0,39; 2,04] 0,7715	-1,99 [-15,50; 11,52] 0,7713
Erkrankungen der Atemwege, des Brustraums und Mediastinums				
7/59 (11,9)	9/58 (15,5)	0,73 [0,25; 2,14] 0,5674	0,76 [0,30; 1,94] 0,5681	-3,65 [-16,23; 8,93] 0,5663

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Erkrankungen des Gastrointestinaltrakts				
25/59 (42,4)	11/58 (19,0)	3,14 [1,35; 7,31] 0,0083	2,23 [1,21; 4,14] 0,0110	23,41 [7,09; 39,73] 0,0053
Erkrankungen der Haut und des Unterhautgewebes				
16/59 (27,1)	15/58 (25,9)	1,07 [0,47; 2,45] 0,8779	1,05 [0,57; 1,93] 0,8779	1,26 [-14,90; 17,42] 0,8779
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen				
26/59 (44,1)	27/58 (46,6)	0,90 [0,43; 1,89] 0,7878	0,95 [0,63; 1,42] 0,7878	-2,48 [-20,71; 15,74] 0,7877
Allgemeine Erkrankungen und Beschwerden am Verabreichungsort				
12/59 (20,3)	6/58 (10,3)	2,21 [0,76; 6,44] 0,1433	1,97 [0,78; 4,93] 0,1483	9,90 [-3,06; 23,05] 0,1322
Untersuchungen				
1/59 (1,7)	6/58 (10,3)	0,15 [0,02; 1,31] 0,0858	0,16 [0,02; 1,35] 0,0919	-8,65 [-17,24; -0,06] 0,0485
Verletzung, Vergiftung und durch Eingriffe bedingte Komplikationen				
16/59 (27,1)	12/58 (20,7)	1,43 [0,60; 3,39] 0,4179	1,31 [0,68; 2,54] 0,4197	6,43 [-9,14; 22,00] 0,4151
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; SOC: Systemorganklasse (System Organ Class); UE: Unerwünschtes Ereignis. Quelle: Sanofi 2024h				

Für UE gesamt zeigte die SOC „Untersuchungen“ bei mindestens 10 % der Patienten in einem Studienarm, ausschließlich für die RD einen statistisch signifikanten Unterschied zugunsten von Sarilumab + CST, der ohne klinische Relevanz war, während sich die SOC „Erkrankungen des Blutes und des Lymphsystems“ und „Erkrankungen des Gastrointestinaltrakts“ statistisch signifikant zuungunsten von Sarilumab + CST gegenüber dem Kontrollarm unterschieden. Für die übrigen SOC ergaben sich keine statistisch signifikanten Unterschiede zwischen den Studienarmen (Tabelle 4-59).

Tabelle 4-60: Ergebnisse der UE gesamt nach PT aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Arthralgie				
9/59 (15,3)	3/58 (5,2)	3,30 [0,83; 13,07] 0,0884	2,95 [0,83; 10,49] 0,0940	10,08 [-0,83; 21,00] 0,0699
Neutropenie				
9/59 (15,3)	0/58 (0,0)	NI		
Diarrhö				
7/59 (11,9)	1/58 (1,7)	7,67 [0,89; 65,96] 0,0632	6,88 [0,86; 55,39] 0,0696	10,14 [1,14; 19,14] 0,0276
Hypertonie				
6/59 (10,2)	2/58 (3,4)	3,17 [0,60; 16,69] 0,1716	2,95 [0,61; 14,25] 0,1765	6,72 [-2,40; 15,85] 0,1473
Schlaflosigkeit				
6/59 (10,2)	9/58 (15,5)	0,62 [0,20; 1,88] 0,3919	0,66 [0,25; 1,74] 0,3937	-5,35 [-17,57; 6,88] 0,3880
Osteoarthritis				
6/59 (10,2)	5/58 (8,6)	1,20 [0,34; 4,23] 0,7749	1,18 [0,38; 3,70] 0,7750	1,55 [-9,13; 12,23] 0,7744
Depression				
5/59 (8,5)	6/58 (10,3)	0,80 [0,23; 2,83] 0,7299	0,82 [0,26; 2,57] 0,7301	-1,87 [-12,56; 8,82] 0,7296
Sturz				
3/59 (5,1)	6/58 (10,3)	0,46 [0,11; 1,98] 0,2973	0,49 [0,13; 1,90] 0,3002	-5,26 [-15,00; 4,48] 0,2869
Nasopharyngitis				
2/59 (3,4)	6/58 (10,3)	0,30 [0,06; 1,60] 0,1585	0,33 [0,07; 1,58] 0,1633	-6,95 [-16,15; 2,24] 0,1367

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; NI: Nicht interpretierbar; OR: Odds Ratio; PT: Bevorzugter Begriff (Preferred Term); RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis. Quelle: Sanofi 2024h				

Für UE gesamt zeigte ausschließlich der PT „Diarrhö“ bei mindestens 10 % der Patienten in einem Studienarm einen statistisch signifikanten, aber klinisch nicht relevanten, Unterschied bei der RD, der zuungunsten von Sarilumab + CST ausfiel. Bei den übrigen PT ergaben sich keine statistisch signifikanten Unterschiede zwischen den Studienarmen. Für den PT „Neutropenie“ war der Unterschied nicht interpretierbar, da im Kontrollarm kein Ereignis auftrat (Tabelle 4-60).

Schwere UE wurden weder auf SOC- noch auf PT-Ebene bei mindestens 5 % der Patienten in einem Studienarm berichtet.

Tabelle 4-61: Ergebnisse der SUE nach SOC aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST		
n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert
Infektionen und parasitäre Erkrankungen				
3/59 (5,1)	3/58 (5,2)	0,98 [0,19; 5,17] 0,9829	0,98 [0,20; 4,75] 0,9829	-0,09 [-8,17; 7,99] 0,9829
Gefäßerkrankungen				
2/59 (3,4)	3/58 (5,2)	0,64 [0,10; 4,08] 0,6369	0,66 [0,11; 3,85] 0,6373	-1,78 [-9,20; 5,63] 0,6348
Skelettmuskulatur-, Bindegewebs- und Knochenkrankungen				
1/59 (1,7)	3/58 (5,2)	0,32 [0,03; 3,21] 0,3270	0,33 [0,03; 3,13] 0,3297	-3,48 [-10,13; 3,18] 0,3027
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; SOC: Systemorganklasse (System Organ Class); SUE: Schwerwiegendes unerwünschtes Ereignis. Quelle: Sanofi 2024h				

Für SUE nach den SOC bei mindestens 5 % der Patienten in einem Studienarm traten bei „Infektionen und parasitäre Erkrankungen“, „Gefäßerkrankungen“ und „Skelettmuskulatur-, Bindegewebs- und Knochenerkrankungen“ keine statistisch signifikanten Unterschiede zwischen den Studienarmen auf (Tabelle 4-61).

SUE traten für keinen PT bei mindestens 5 % der Patienten in einem Studienarm auf.

Tabelle 4-62: Ergebnisse für UE, die zum Therapieabbruch führten, nach SOC und PT aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

SOC, n (%) PT, n (%)	Sarilumab + CST (N=59)	Placebo + CST (N=58)
Anzahl aller Patienten mit Ereignis, n (%)	7 (11,9)	4 (6,9)
Infektionen und parasitäre Erkrankungen	3 (5,1)	1 (1,7)
COVID-19	1 (1,7)	1 (1,7)
Bandscheibenvorfall	1 (1,7)	0
Pneumonie	1 (1,7)	0
Erkrankungen des Blutes und des Lymphsystems	3 (5,1)	0
Neutropenie	3 (5,1)	0
Gefäßerkrankungen	0	1 (1,7)
Riesenzellarteriitis	0	1 (1,7)
Erkrankungen der Atemwege, des Brustraums und Mediastinums	0	1 (1,7)
Lungenembolie	0	1 (1,7)
Erkrankungen der Nieren und Harnwege	1 (1,7)	0
Nierensteine	1 (1,7)	0
Untersuchungen	0	1 (1,7)
ALT erhöht	0	1 (1,7)
ALT: Alanin-Aminotransferase; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); N: Anzahl der Patienten; n: Anzahl der Patienten mit mindestens einem Ereignis; PT: Bevorzugter Begriff (Preferred Term); RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SOC: Systemorganklasse (System Organ Class); UE: Unerwünschtes Ereignis. Quelle: Sanofi 2024h		

Auch die Anzahl der Patienten mit UE, die zum Therapieabbruch führten, fiel mit $n = 7$ (11,9 %) im Sarilumab-Arm bzw. $n = 4$ (6,9 %) im Kontrollarm niedrig aus. Im Sarilumab-Arm waren die SOC „Infektionen und parasitäre Erkrankungen“ und SOC „Erkrankungen des Blutes und des Lymphsystems“ mit jeweils 3 Patienten (5,1 %) am häufigsten betroffen. In der letztgenannten SOC wurde „Neutropenie“ bei 3 Patienten (5,1 %) als häufigster Grund des Therapieabbruchs angegeben. Im Kontrollarm führte kein UE in einer SOC bei mehr als einem Patienten (1,7 %) zum Therapieabbruch (Tabelle 4-62).

Tabelle 4-63: Ergebnisse für UE von besonderem Interesse mit PT aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

UE von besonderem Interesse, n (%) PT, n (%)	Sarilumab + CST (N=59)	Placebo + CST (N=58)
Anzahl aller Patienten mit Ereignis, n (%)	7 (11,9)	5 (8,6)
Opportunistische Infektionen^a	2 (3,4)	2 (3,4)
Herpes Zoster	2 (3,4)	1 (1,7)
COVID-19	0	1 (1,7)
Infektionen mit längerfristigem Medikationsbedarf	2 (3,4)	2 (3,4)
Bakterielle Zystitis	1 (1,7)	0
Bandscheibenvorfall	1 (1,7)	0
COVID-19	0	1 (1,7)
Wundrose	0	1 (1,7)
Infektionen mit Bedarf an parenteraler Ernährung	1 (1,7)	2 (3,4)
Bandscheibenvorfall	1 (1,7)	0
COVID-19	0	1 (1,7)
COVID-19-Pneumonie	0	1 (1,7)
ALT-Anstieg auf ≥ 3 x ULN	0	2 (3,4)
ALT erhöht	0	2 (3,4)
ALT-Anstieg, der zum Therapieabbruch führte	0	1 (1,7)
ALT erhöht	0	1 (1,7)
ANC-Abnahme, die zum Therapieabbruch führte	3 (5,1)	0
Neutropenie	3 (5,1)	0
a: Prüfarztbeurteilung. ALT: Alanin-Aminotransferase; ANC: Absolute Neutrophilenzahl (Absolute Neutrophil Count); CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); N: Anzahl der Patienten; n: Anzahl der Patienten mit mindestens einem Ereignis; PT: Bevorzugter Begriff (Preferred Term); RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); UE: Unerwünschtes Ereignis; ULN: Obere Normgrenze (Upper Limit of Normal). Quelle: Sanofi 2021a		

Auf eine Responderanalyse wird für die einzelnen UE von besonderem Interesse aufgrund der geringen Anzahl an Patienten verzichtet. Insgesamt traten bei 7 Patienten (11,9 %) im Sarilumab-Arm und fünf Patienten (8,6 %) im Kontrollarm UE von besonderem Interesse auf. Am häufigsten vertreten war „Neutropenie“, die zum Therapieabbruch führte, bei drei Patienten im Sarilumab-Arm und erhöhte ALT (≥ 3 x ULN) bei zwei Patienten im Kontrollarm. Die übrigen UE von besonderem Interesse, „Schwangerschaft einer Studienpatientin oder Partnerin eines Studienpatienten“, „Symptomatische Überdosierung der Studienmedikation“ sowie „Thrombozytopenie“ wurden im Studienverlauf bei keinem Patienten beobachtet (Tabelle 4-63).

4.3.1.3.2 Subgruppenanalysen – RCT

Für die Darstellung der Ergebnisse aus Subgruppenanalysen gelten die gleichen Anforderungen wie für die Darstellung von Ergebnissen aus Gesamtpopulationen in Abschnitt 4.3.1.3.1.¹⁶

Darüber hinaus sind folgende Kriterien zu berücksichtigen:

- Subgruppenanalysen sind nur für die Merkmale (z.B. Alter) durchzuführen, bei denen die resultierenden Subgruppen jeweils mindestens 10 Patienten umfassen.
- Subgruppenanalysen sind für binäre Ereignisse je Merkmal nur dann durchzuführen, wenn in einer der Subgruppen mindestens 10 Ereignisse aufgetreten sind.
- Für Überlebenszeitanalysen müssen Kaplan-Meier-Kurven zu den einzelnen Subgruppen nur für Subgruppenanalysen mit statistisch signifikantem Interaktionsterm ($p < 0,05$) dargestellt werden.
- Ergebnisse zu UE nach SOC und PT müssen nur dargestellt werden, wenn das jeweilige Ergebnis für die Gesamtpopulation statistisch signifikant ist. Zu a priori definierten Ereignissen (z.B. AESI, SMQs) sowie den UE-Gesamtraten (UE, schwere UE, SUE und Abbrüche wegen UE) müssen Subgruppenanalysen unabhängig vom Vorliegen statistischer Signifikanz in der Gesamtpopulation dargestellt werden.
- Bei Vorliegen mehrerer Studien und Durchführung von Metaanalysen zu diesen Studien gelten die zuvor genannten Kriterien für die jeweilige Metaanalyse, nicht für die Einzelstudien.
- Für Studien des pharmazeutischen Unternehmers sind entsprechende Analysen für alle benannten Effektmotifikatoren zu allen relevanten Endpunkten nach den zuvor genannten Kriterien vorzulegen und daher ggf. posthoc durchzuführen.
- Wird für die Nutzenbewertung nur die Teilpopulation einer Studie herangezogen (z.B. wegen Zulassungsbeschränkungen, aufgrund von durch den G-BA bestimmte Teilpopulationen), so gelten die genannten Kriterien für diese Teilpopulation, und die Subgruppenanalysen sind für die Teilpopulation und nicht für die Gesamtpopulation der Studie durchzuführen.
- Subgruppenanalysen, bei denen der Interaktionsterm nicht statistisch signifikant ist, können auch in einem separaten Anhang des vorliegenden Modul 4 dargestellt werden. Dabei kann die Ausgabe der Statistik-Software unverändert verwendet werden, sofern diese alle notwendigen Angaben enthält. Eine ausschließliche Darstellung in Modul 5 ist aber nicht ausreichend.

¹⁶ unbesetzt

Beschreiben Sie die Ergebnisse von Subgruppenanalysen. Stellen Sie dabei zunächst tabellarisch dar, zu welchen der in Abschnitt 4.2.5.5 genannten Effektmodifikatoren Subgruppenanalysen zu den relevanten Endpunkten vorliegen, und ob diese a priori geplant und im Studienprotokoll festgelegt waren oder posthoc durchgeführt wurden.

Orientieren Sie sich an der beispielhaften Angabe in der ersten Tabellenzeile.

Tabelle 4-64: Matrix der durchgeführten Subgruppenanalysen (Studie SAPHYR)

Altersgruppe	Geschlecht	Abstammung	Geografische Region ^a	CRP zu Baseline	ESR zu Baseline	CS-Dosis zu Baseline	Anzahl Rezidive vor Screening	MTX Add-on-Therapie
Anhaltende Remission zu Woche 52^b								
●	●	●	○	●	●	●	●	○
Zeit bis zum ersten PMR-Schub nach klinischer Remission								
○	○	○	○	○	○	○	○	○
PMR-AS^c								
○	○	○	○	○	○	○	○	○
FACIT-F								
○	○	○	○	○	○	○	○	○
EQ-5D VAS								
○	○	○	○	○	○	○	○	○
HAQ-DI								
○	○	○	○	○	○	○	○	○
SF-36v2 - PCS								
○	○	○	○	○	○	○	○	○
SF-36v2 - MCS								
○	○	○	○	○	○	○	○	○
Nebenwirkungen - UE gesamt								
○	○	○	○	○	○	○	○	○
Nebenwirkungen - schwere UE								
○	○	○	○	○	○	○	○	○
Nebenwirkungen – SUE								
○	○	○	○	○	○	○	○	○
Nebenwirkungen – UE, die zum Therapieabbruch führten								
○	○	○	○	○	○	○	○	○

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Altersgruppe	Geschlecht	Abstammung	Geografische Region ^a	CRP zu Baseline	ESR zu Baseline	CS-Dosis zu Baseline	Anzahl Rezidive vor Screening	MTX Add-on-Therapie
<p>a: A priori unterteilt in 3 geografische Regionen: westliche Länder, Südamerika und Rest der Welt. Da insgesamt weniger als 10 südamerikanische Patienten an der Studie teilnahmen, wurden die Regionen Südamerika und Rest der Welt zu einer Subgruppe zusammengefasst. Alle Subgruppenanalysen erfolgten für die beiden geografischen Regionen Westliche Länder und Rest der Welt post-hoc.</p> <p>b: einschließlich der Einzelkomponenten „Remission spätestens zu Woche 12“, „kein Krankheitsschub von Woche 12 bis 52“, „anhaltende CRP-Reduktion von Woche 12 bis 52“, „erfolgreiches Prednison-Ausschleichen von Woche 12 bis Woche 52“.</p> <p>c: einschließlich der Einzelkomponenten „CRP“, „VAS (Patienten- und Prüfarztbeurteilung)“, „Dauer der Morgensteifigkeit“, „Beweglichkeit der oberen Gliedmaßen“.</p> <p>●: A priori geplante Subgruppenanalyse. ○: Post hoc durchgeführte Subgruppenanalyse</p> <p>BMI: Body Mass Index; CRP: C-reaktives Protein; CS: Corticosteroids; EQ-5D: European Quality of Life 5 Dimensions; ESR: Erythrozyten-Sedimentationsrate; FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI: Health Assessment Questionnaire Disability Index; MCS: Summenskala der psychischen Gesundheit (Mental Health Component Summary Score); MTX: Methotrexat; PCS: Summenskala der körperlichen Gesundheit (Physical Health Component Summary Score); PMR-AS: Polymyalgia rheumatica-Aktivitätsscore; SF-36v2: 36-Item Short Form Health Survey Version 2; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala.</p> <p>Quellen: Sanofi 2021d; Sanofi 2024b, 2024d, 2024f, 2024i</p>								

Stellen Sie anschließend in Tabelle 4-65 die Ergebnisse der Interaktionsterme für alle Subgruppenanalysen je Endpunkt in tabellarischer Form dar, und zwar für jede einzelne Studie separat. Kennzeichnen Sie dabei statistisch signifikante ($p < 0,05$) Interaktionsterme.

Tabelle 4-65: Interaktionsterme (p-Werte) der Subgruppenanalysen je Endpunkt für Studie SAPHYR und Subgruppenmerkmale

Altersgruppe	Geschlecht	Abstammung	Geografische Region	CRP zu Baseline	ESR zu Baseline	CS-Dosis zu Baseline	Anzahl Rezidive vor Screening	MTX Add-on-Therapie
Anhaltende Remission zu Woche 52								
0,0973	0,5017	0,9253	k.A.	k.A.	k.A.	0,9748	0,2641	0,7292
Anhaltende Remission zu Woche 52 – Remission spätestens zu Woche 12								
0,5060	0,4635	0,9730	0,4881	0,1742	0,5895	0,0100	0,7358	0,9182
Anhaltende Remission zu Woche 52 – kein Krankheitsschub von Woche 12 bis 52								
0,0887	0,8460	0,5200	0,1750	0,9050	0,0332	0,2197	0,7293	0,9533
Anhaltende Remission zu Woche 52 – anhaltende CRP-Reduktion von Woche 12 bis 52								
0,1816	0,0210	0,8605	0,0412	0,5162	0,0391	0,9514	0,4397	0,3390

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Altersgruppe	Geschlecht	Abstammung	Geografische Region	CRP zu Baseline	ESR zu Baseline	CS-Dosis zu Baseline	Anzahl Rezidive vor Screening	MTX Add-on-Therapie
Anhaltende Remission zu Woche 52 – erfolgreiches Prednison-Ausschleichen von Woche 12 bis 52								
0,0274	0,8479	0,6797	0,1568	0,9813	0,0766	0,6124	0,1257	0,9531
Zeit bis zum ersten PMR-Schub nach klinischer Remission								
0,0298	0,2573	0,3587	0,7922	0,6627	0,3283	0,9358	0,9426	0,4311
PMR-AS								
0,0553	0,1557	0,1070	0,0505	0,3261	0,6890	0,6422	0,6471	0,7515
PMR-AS - CRP								
0,6392	0,0556	0,8209	0,8852	0,8838	0,6160	0,3180	0,4441	0,7264
PMR-AS – VAS für Schmerz, Patientenbeurteilung								
0,1063	0,7736	0,9652	0,2843	0,0808	0,1960	0,9127	0,2339	0,5677
PMR-AS – VAS für Krankheitsaktivität, Prüfarztbeurteilung								
0,0424	0,1989	0,2685	0,0615	0,9216	0,3991	0,6574	0,0748	0,4304
PMR-AS – Dauer der Morgensteifigkeit								
0,7308	0,8893	0,0125	0,1710	0,1350	0,0638	0,6826	0,4641	0,8476
PMR-AS – Beweglichkeit der oberen Gliedmaßen								
0,1192	0,9939	0,3425	0,5402	0,0131	0,8363	0,7821	0,0851	0,7308
FACIT-F								
0,0498	0,6452	0,4306	0,6880	0,1250	0,4522	0,0241	0,3655	0,7177
EQ-5D VAS								
0,2492	0,3595	0,2456	k.A.	k.A.	k.A.	k.A.	k.A.	0,8538
HAQ-DI – Standard-Score								
0,1615	0,6119	0,6392	0,9549	0,1413	0,3609	0,1302	0,0337	0,8044
HAQ-DI – Globale Krankheitsaktivität								
0,2400	0,2353	0,7641	0,0190	0,4032	0,1549	0,9177	0,3172	0,5691
SF-36v2 - PCS								
k.A.	0,3115	0,1460	0,3895	k.A.	0,3264	k.A.	k.A.	0,9210
SF-36v2 - MCS								
k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.
Nebenwirkungen - UE gesamt								
0,9397	<0,0001	k.A.	<0,0001	<0,0001	<0,0001	0,2975	0,7085	0,4718
Nebenwirkungen – schwere UE								
k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.

Altersgruppe	Geschlecht	Abstammung	Geografische Region	CRP zu Baseline	ESR zu Baseline	CS-Dosis zu Baseline	Anzahl Rezidive vor Screening	MTX Add-on-Therapie
Nebenwirkungen – SUE								
k.A.	k.A.	k.A.	0,9739	k.A.	0,0258	k.A.	k.A.	k.A.
Nebenwirkungen – UE, die zum Therapieabbruch führten								
k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.	k.A.
<p>CRP: C-reaktives Protein; CS: Corticosteroids; EQ-5D: European Quality of Life 5 Dimensions; ESR: Erythrozyten-Sedimentationsrate; FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI: Health Assessment Questionnaire Disability Index; k.A.: Keine Angabe; MCS: Summenskala der psychischen Gesundheit (Mental Health Component Summary Score); MTX: Methotrexat; PCS: Summenskala der körperlichen Gesundheit (Physical Health Component Summary Score); PMR-AS: Polymyalgia rheumatica-Aktivitätsscore; SF-36v2: 36-Item Short Form Health Survey Version 2; SUE: Schwerwiegendes unerwünschtes Ereignis; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala.</p> <p>Quellen: Sanofi 2024b, 2024d, 2024f, 2024i</p>								

Stellen Sie schließlich alle Subgruppenergebnisse dar.

Sofern eine Effektmodifikation für mehr als ein Subgruppenmerkmal vorliegt, kann eine Untersuchung auf eine Wechselwirkung höherer Ordnung sinnvoll sein. Dies gilt insbesondere dann, wenn diese Effektmodifikation konsistent über mehrere Endpunkte besteht. Zur Interpretation der Ergebnisse sollte dann für diese Endpunkte zusätzlich eine Subgruppenanalyse durchgeführt werden, die die Merkmale mit Effektmodifikation kombiniert. Beispiel: Für die Endpunkte Mortalität, gesundheitsbezogene Lebensqualität und schwere unerwünschte Ereignisse liegt sowohl für das Merkmal Geschlecht (mit den Ausprägungen „weiblich“ und „männlich“) als auch für das Merkmal Schweregrad (mit den Ausprägungen „niedrig“ und „hoch“) eine Effektmodifikation vor. Die zusätzliche Subgruppenanalyse erfolgt dann für die 3 genannten Endpunkte für das kombinierte Merkmal Geschlecht/Schweregrad mit den 4 Ausprägungen weiblich/niedrig, weiblich/hoch, männlich/niedrig und männlich/hoch.

Sofern die vorliegenden Studien bzw. Daten für eine Meta-Analyse medizinisch und methodisch geeignet sind, fassen Sie die Ergebnisse mithilfe einer Meta-Analyse quantitativ zusammen und stellen Sie die Ergebnisse der Meta-Analyse (als Forest-Plot) dar.

Beschreiben Sie die Ergebnisse zusammenfassend. Begründen Sie Ihr Vorgehen, wenn Sie keine Meta-Analyse durchführen bzw. wenn Sie nicht alle Studien in die Meta-Analyse einschließen.

Im Folgenden werden die Subgruppenanalysen mit signifikantem Interaktions-p-Wert je Endpunkt dargestellt. Eine endpunktübergreifende Effektmodifikation konnte nicht festgestellt werden.

4.3.1.3.2.1 Subgruppenanalysen für den Endpunkt Anhaltende Remission zu Woche 52 – RCT

Tabelle 4-66: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Remission spätestens zu Woche 12 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions-test
Alle Patienten	28/60 (46,7)	22/58 (37,9)	1,43 [0,68; 3,01] 0,3397	1,23 [0,80; 1,89] 0,3454	8,74 [-9,21; 26,68] 0,3369	Nicht zutreffend
CS-Dosis zu Baseline (mg)						
≤ Median	13/30 (43,3)	19/34 (55,9)	0,60 [0,22; 1,66] 0,3215	0,78 [0,46; 1,30] 0,3347	-12,55 [-37,38; 12,29] 0,3164	0,0100
> Median	15/30 (50,0)	3/24 (12,5)	7,00 [1,66; 29,52] 0,0090	4,00 [1,27; 12,56] 0,0175	37,50 [14,72; 60,28] 0,0017	
CS: Corticosteroids; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko. Quellen: Sanofi 2024d, 2024e						

Für die Einzelkomponente Remission spätestens zu Woche 12 des Endpunkts Anhaltende Remission zu Woche 52 ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „CS-Dosis zu Baseline (mg)“ ($p=0,0100$). Der Effektschätzer in der Subgruppenausprägung „≤ Median“ ist nicht gleichgerichtet zum Gesamteffekt. Das Ergebnis ist jedoch nicht statistisch signifikant. Für die Subgruppenausprägung „> Median“ liegt ein statistisch signifikanter Vorteil für Sarilumab vor (Tabelle 4-66). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Tabelle 4-67: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Kein Krankheitsschub von Woche 12 bis 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions-test
Alle Patienten	33/60 (55,0)	19/58 (32,8)	2,51 [1,18; 5,34] 0,0175	1,68 [1,08; 2,60] 0,0206	22,24 [4,61; 39,87] 0,0139	Nicht zutreffend
ESR zu Baseline (mm/h)						
≤ Median	18/28 (64,3)	7/31 (22,6)	6,17 [1,92; 19,84] 0,0028	2,85 [1,38; 5,87] 0,0046	41,71 [18,15; 65,26] 0,0008	0,0332
> Median	15/32 (46,9)	12/27 (44,4)	1,10 [0,39; 3,16] 0,8526	1,05 [0,60; 1,87] 0,8553	2,43 [-23,62; 28,48] 0,8525	
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ESR: Erythrozyten-Sedimentationsrate; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko. Quellen: Sanofi 2024d, 2024e						

Für die Einzelkomponente Kein Krankheitsschub von Woche 12 bis 52 des Endpunkts Anhaltende Remission zu Woche 52 ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „ESR zu Baseline (mm/h)“ ($p=0,0332$). Die Effektschätzer in den Subgruppenausprägungen „≤ Median“ und „> Median“ sind gleichgerichtet zum Gesamteffekt (Tabelle 4-67). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Tabelle 4-68: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Anhaltende CRP-Reduktion von Woche 12 bis 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions- test
Alle Patienten	40/60 (66,7)	26/58 (44,8)	2,46 [1,16; 5,23] 0,0195	1,49 [1,06; 2,09] 0,0223	21,84 [4,16; 39,52] 0,0159	Nicht zutreffend
Geschlecht						
Männlich	9/15 (60,0)	14/21 (66,7)	0,75 [0,18; 3,12] 0,6843	0,90 [0,53; 1,53] 0,6973	-6,67 [-39,80; 26,47] 0,6852	0,0210
Weiblich	31/45 (68,9)	12/37 (32,4)	4,61 [1,79; 11,91] 0,0019	2,12 [1,27; 3,55] 0,0040	36,46 [15,89; 57,03] 0,0007	
Geografische Region						
Westliche Länder	30/47 (63,8)	15/44 (34,1)	3,41 [1,42; 8,17] 0,0064	1,87 [1,17; 3,00] 0,0089	29,74 [9,85; 49,63] 0,0038	0,0412
Rest der Welt	10/13 (76,9)	11/14 (78,6)	0,91 [0,13; 6,12] 0,9188	0,98 [0,64; 1,50] 0,9221	-1,65 [-34,65; 31,36] 0,9189	
ESR zu Baseline (mm/h)						
≤ Median	20/28 (71,4)	10/31 (32,3)	5,25 [1,68; 16,37] 0,0050	2,21 [1,25; 3,93] 0,0066	39,17 [15,19; 63,15] 0,0018	0,0391
> Median	20/32 (62,5)	16/27 (59,3)	1,15 [0,39; 3,35] 0,8002	1,05 [0,69; 1,61] 0,8042	3,24 [-22,30; 28,78] 0,8003	
CRP: C-reaktives Protein; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ESR: Erythrozyten-Sedimentationsrate; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko.						
Quellen: Sanofi 2024d, 2024e						

Für die Einzelkomponente Anhaltende CRP-Reduktion von Woche 12 bis 52 des Endpunkts Anhaltende Remission zu Woche 52 ergaben sich statistisch signifikante Interaktionen zwischen der Behandlung und den subgruppenbildenden Merkmalen „Geschlecht“ (p=0,0210), „Geografische Region“ (p=0,0412) und „ESR zu Baseline (mm/h)“ (p=0,0391). Für das subgruppenbildende Merkmal „Geschlecht“ zeigte sich in der Subgruppenausprägung

„Männlich“ und für das subgruppenbildende Merkmal „Geografische Region“ in der Subgruppenausprägung „Rest der Welt“ gegenüber der Gesamtpopulation ein Subgruppeneffekt. Es liegen qualitative Unterschiede vor. Die Effektschätzer in den Subgruppenausprägungen „Männlich“ und „Rest der Welt“ sind nicht gleichgerichtet zum Gesamteffekt. Das 95 %-KI des jeweiligen RR schließt jedoch die 1 mit ein. Für das subgruppenbildende Merkmal „ESR zu Baseline (mm/h)“ sind die Effektschätzer in den Subgruppenausprägungen „≤ Median“ und „> Median“ gleichgerichtet zum Gesamteffekt (Tabelle 4-68). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Tabelle 4-69: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Erfolgreiches Prednison-Ausschleichen von Woche 12 bis 52 aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions- test
Alle Patienten	30/60 (50,0)	14/58 (24,1)	3,14 [1,42; 6,95] 0,0051	2,07 [1,22; 3,51] 0,0068	25,86 [8,91; 42,81] 0,0031	Nicht zutreffend
Altersgruppe (Jahre)						
< 70	22/34 (64,7)	4/25 (16,0)	9,62 [2,60; 35,59] 0,0010	4,04 [1,56; 10,48] 0,0040	48,71 [26,69; 70,73] < 0,0001	0,0274
≥ 70	8/26 (30,8)	10/33 (30,3)	1,02 [0,33; 3,20] 0,9693	1,02 [0,46; 2,24] 0,9698	0,47 [-23,72; 24,66] 0,9694	
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko.						
Quellen: Sanofi 2024d, 2024e						

Für die Einzelkomponente Erfolgreiches Prednison-Ausschleichen von Woche 12 bis 52 des Endpunkts Anhaltende Remission zu Woche 52 ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „Altersgruppe (Jahre)“ ($p=0,0274$). Die Effektschätzer in den Subgruppenausprägungen „< 70“ und „≥ 70“ sind gleichgerichtet zum Gesamteffekt (Tabelle 4-69). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante

Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

4.3.1.3.2 Subgruppenanalysen für den Endpunkt Kumulative Corticosteroid-Gesamtdosis über 52 Wochen – RCT

Das im Verlauf der Studie SAPHYR geplante Prednison-Ausschleichen resultierte in einem deutlichen Unterschied sowohl der zu erwartenden als auch der tatsächlichen CS-Gesamtdosis in den beiden Studienarmen. Da die Effekte der Subgruppenanalysen gleichgerichtet sind zum Ergebnis der Gesamtpopulation und der Endpunkt nicht für die Ableitung eines Zusatznutzens herangezogen wird, wird an dieser Stelle auf die Darstellung der Subgruppenanalysen verzichtet. Die Ergebnisse befinden sich in Anhang 4-G.

4.3.1.3.2.3 Subgruppenanalysen für den Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission – RCT

Tabelle 4-70: Ergebnisse der Subgruppenanalysen für die Zeit bis zum ersten PMR-Schub nach klinischer Remission aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST		Placebo + CST		Sarilumab + CST versus Placebo + CST	
	n/N (%)	Median (Monate) [95 %-KI] ^a	n/N (%)	Median (Monate) [95 %-KI] ^a	HR [95 %-KI] ^b p-Wert ^b	p-Wert für Interaktions-test ^c
Alle Patienten	29/60 (48,3)	NE [93,0; NE]	45/58 (77,6)	99,0 [1,0; 154,0]	0,56 [0,35; 0,90] 0,0153	Nicht zutreffend
Altersgruppe (Jahre)						
< 70	12/34 (35,3)	NE [14,0; NE]	21/25 (84,0)	0,0 [0,0; 20,3]	0,34 [0,17; 0,70] 0,0035	0,0298
≥ 70	17/26 (65,4)	15,9 [0,0; NE]	24/33 (72,7)	19,6 [0,0; 40,0]	0,94 [0,50; 1,75] 0,8443	
<p>a: Kaplan-Meier-Schätzer. b: Cox-Proportional-Hazard-Modell. c: Cox-Proportional-Hazard-Modell mit Behandlung, Subgruppe, Behandlung-nach-Subgruppe als Kovariaten. CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); HR: Hazard Ratio; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; NE: Nicht berechenbar (Not Evaluable); PMR: Polymyalgia rheumatica; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial). Quellen: Sanofi 2024f, 2024g</p>						

Für den Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „Altersgruppe (Jahre)“ (p=0,0298). Die Effektschätzer in den

Subgruppenausprägungen „< 70“ und „≥ 70“ sind gleichgerichtet zum Gesamteffekt (Tabelle 4-70). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Die Kaplan-Meier-Kurve für den Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission zur Subgruppe „Altersgruppe (Jahre)“ ist in Abbildung 4-15 dargestellt.

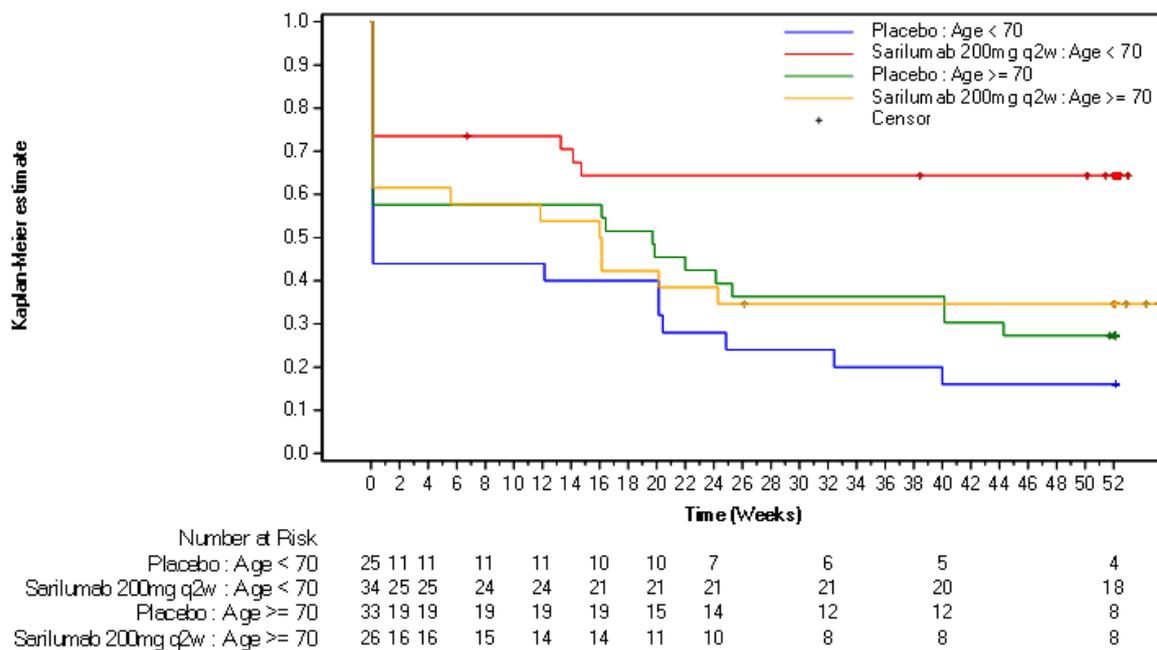


Abbildung 4-15: Kaplan-Meier-Kurve für den Endpunkt Zeit bis zum ersten PMR-Schub nach klinischer Remission, Subgruppenmerkmal „Altersgruppe (Jahre)“, aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

ITT: Intention-To-Treat; PMR: Polymyalgia rheumatica; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial).

Quelle: Sanofi 2024f

4.3.1.3.2.4 Subgruppenanalysen für den Endpunkt PMR-AS – RCT

Tabelle 4-71: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Veränderung der VAS um $\geq -1,5$ Punkte für Krankheitsaktivität, Prüfarztbeurteilung, aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions-test
Alle Patienten	30/60 (50,0)	21/58 (36,2)	1,76 [0,84; 3,71] 0,1346	1,38 [0,90; 2,12] 0,1408	13,79 [-4,09; 31,67] 0,1292	Nicht zutreffend
Altersgruppe (Jahre)						
< 70	21/34 (61,8)	7/25 (28,0)	4,15 [1,33; 12,96] 0,0151	2,21 [1,10; 4,43] 0,0260	33,76 [9,23; 58,30] 0,0078	0,0424
≥ 70	9/26 (34,6)	14/33 (42,4)	0,72 [0,24; 2,13] 0,5445	0,82 [0,42; 1,60] 0,5550	-7,81 [-33,22; 17,61] 0,5408	
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; PMR-AS: Polymyalgia rheumatica-Aktivitätsscore; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; VAS: Visuelle Analogskala.						
Quellen: Sanofi 2024d, 2024e						

Für die Einzelkomponente Veränderung der VAS für Krankheitsaktivität, Prüfarztbeurteilung, um $\geq -1,5$ Punkte des Endpunkts PMR-AS ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „Altersgruppe (Jahre)“ ($p=0,0424$). Für die Subgruppenausprägung „ ≥ 70 “ zeigte sich gegenüber der Gesamtpopulation ein Subgruppeneffekt. Es liegt ein qualitativer Unterschied vor. Der Effektschätzer in der Subgruppenausprägung „ ≥ 70 “ ist nicht gleichgerichtet zum Gesamteffekt. Das 95 %-KI des RR schließt jedoch die 1 mit ein. Für die Subgruppenausprägung „ < 70 “ zeigte sich ebenfalls gegenüber der Gesamtpopulation ein Subgruppeneffekt. Es liegt ein statistisch signifikanter Vorteil für Sarilumab vor (Tabelle 4-71). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Für die Einzelkomponente Veränderung der VAS für Schmerz, Patientenbeurteilung, um $\geq -1,5$ Punkte des Endpunkts PMR-AS ergaben sich zwischen der Behandlung und den subgruppenbildenden Merkmalen für keine Subgruppe statistisch signifikante Interaktionen. Die Ergebnisse der Subgruppenanalysen befinden sich ebenfalls in Anhang 4-G.

Tabelle 4-72: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Dauer der Morgensteifigkeit aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

Sarilumab + CST (N=60)			Placebo + CST (N=58)			Sarilumab + CST versus Placebo + CST		
n/ N	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	n/ N	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	LS-MWD [95 %-KI] ^a p-Wert ^a Hedges' g [95 %-KI]	p-Wert für Inter- aktionstest ^b	
Alle Patienten								
48/ 60	66,35 (64,86)	-75,61 (5,87)	46/ 58	106,30 (216,84)	-53,18 (5,98)	-22,43 [-39,09; -5,77] 0,0088 -0,55 [-0,96; -0,14]	Nicht zutreffend	
Abstammung (weiß)								
41/ 50	68,17 (59,12)	-77,09 (5,28)	38/ 48	112,89 (232,50)	-64,34 (5,46)	-12,75 [-27,91; 2,41] 0,0982 -0,37 [-0,92; 0,07]	0,0125	
Abstammung (andere)								
7/ 10	55,71 (97,49)	NI (NI)	8/ 10	75,00 (123,17)	NI (NI)	NI NI		
Fehlende Werte wurden nicht imputiert.								
a: MMRM mit Veränderung gegenüber Baseline als abhängige Variable und Baseline-Wert, Behandlungsgruppe, Visite und Interaktionen zwischen Baseline-Wert und Visite sowie zwischen Behandlung und Visite als Kovariaten.								
b: MMRM mit Veränderung gegenüber Baseline als abhängige Variable und Baseline-Wert, Behandlungsgruppe, Visite und Interaktionen zwischen Baseline-Wert und Visite, Behandlung und Visite, Behandlung und Subgruppe sowie zwischen Behandlung und Subgruppe und Visite als Kovariaten.								
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall; LS-MW: Kleinste-Quadrat-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrat-Mittelwertdifferenz (Least Square-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit einer Erhebung zu Baseline und zu Woche 52; NI: Nicht interpretierbar; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error).								
Quelle: Sanofi 2024b, 2024c								

Für die Einzelkomponente Dauer der Morgensteifigkeit des Endpunkts PMR-AS ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „Abstammung“ (p=0,0125). Die Ergebnisse der Subgruppenausprägung „andere“ ist aufgrund der geringen Fallzahl in diesem Stratum nicht interpretierbar. Potenzielle Effektunterschiede können somit nicht beurteilt werden. Für die

übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Das Longitudinalplot für die Einzelkomponente Dauer der Morgensteifigkeit zur Subgruppe „Abstammung“ ist in Abbildung 4-16 dargestellt.

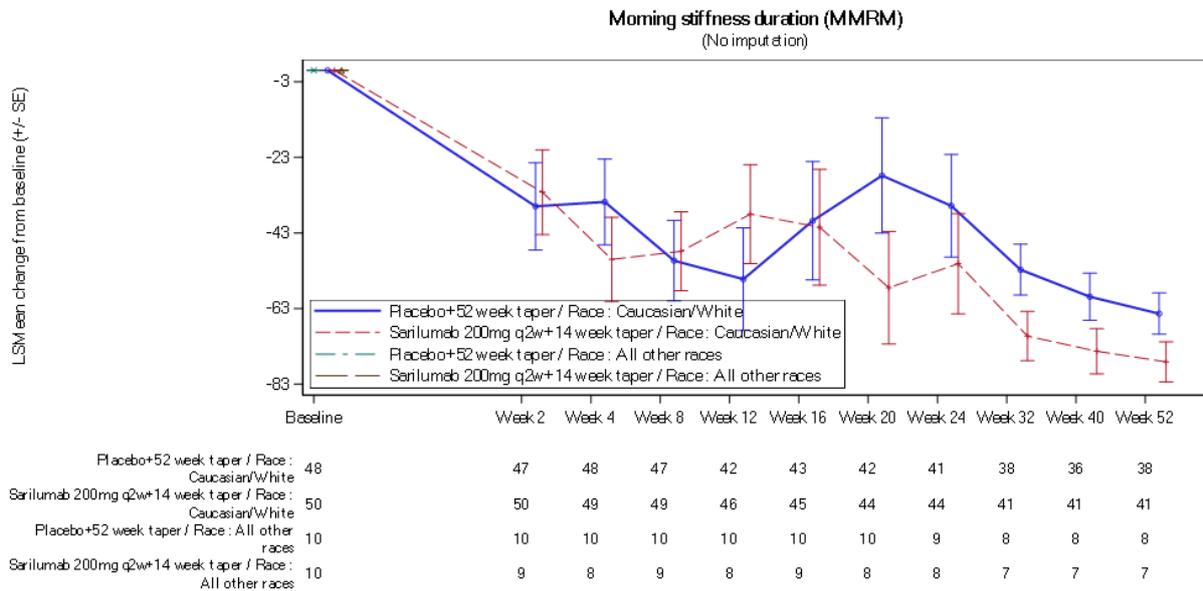


Abbildung 4-16: Longitudinalplot für die Veränderung der Einzelkomponente Dauer der Morgensteifigkeit, Subgruppenmerkmal „Abstammung“, aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

ITT: Intention-To-Treat; LSMean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error)

Quelle: Sanofi 2024b

Tabelle 4-73: Ergebnisse der Subgruppenanalysen für die Einzelkomponente Beweglichkeit der oberen Gliedmaßen aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel

Sarilumab + CST (N=60)			Placebo + CST (N=58)			Sarilumab + CST versus Placebo + CST	
n/ N	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	n/ N	MW (SD) zu Baseline	LS-MW (SE) ^a Veränderung zu Woche 52	LS-MWD [95 %-KI] ^a p-Wert ^a Hedges' g [95 %-KI]	p-Wert für Inter- aktionstest ^b
Alle Patienten							
48/ 60	0,52 (0,80)	-0,47 (0,06)	46/ 58	0,46 (0,62)	-0,23 (0,06)	-0,24 [-0,40; -0,08] 0,0044 -0,60 [-1,00; -0,19]	Nicht zutreffend
CRP zu Baseline (mg/l) - ≤ Median							
22/ 28	0,45 (0,67)	-0,49 (0,12)	24/ 31	0,58 (0,72)	0,06 (0,11)	-0,55 [-0,87; -0,23] 0,0011 -1,01 [-1,60; -0,42]	0,0131
CRP zu Baseline (mg/l) - > Median							
26/ 32	0,58 (0,90)	-0,43 (0,06)	22/ 27	0,32 (0,48)	-0,42 (0,06)	-0,01 [-0,18; 0,15] 0,8810 -0,04 [-0,62; 0,53]	
Fehlende Werte wurden nicht imputiert.							
a: MMRM mit Veränderung gegenüber Baseline als abhängige Variable und Baseline-Wert, Behandlungsgruppe, Visite und Interaktionen zwischen Baseline-Wert und Visite sowie zwischen Behandlung und Visite als Kovariaten.							
b: MMRM mit Veränderung gegenüber Baseline als abhängige Variable und Baseline-Wert, Behandlungsgruppe, Visite und Interaktionen zwischen Baseline-Wert und Visite, Behandlung und Visite, Behandlung und Subgruppe sowie zwischen Behandlung und Subgruppe und Visite als Kovariaten.							
CRP: C-reaktives Protein; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall; LS-MW: Kleinste-Quadrate-Mittelwert (Least Squares-Mean); LS-MWD: Kleinste-Quadrate-Mittelwertdifferenz (Least Square-Mean Value Difference); MMRM: Mixed Model for Repeated Measures; MW: Mittelwert; N: Anzahl der Patienten; n: Anzahl der Patienten mit einer Erhebung zu Baseline und zu Woche 52; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SD: Standardabweichung (Standard Deviation); SE: Standardfehler (Standard Error)							
Quelle: Sanofi 2024b, 2024c							

Für die Einzelkomponente Beweglichkeit der oberen Gliedmaßen des Endpunkts PMR-AS ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „CRP zu Baseline (mg/l)“ (p=0,0131). Die Effektschätzer in

den Subgruppenausprägungen „ \leq Median“ und „ $>$ Median“ sind gleichgerichtet zum Gesamteffekt. Dabei fällt das Ergebnis für „ \leq Median“ im Gegensatz zu „ $>$ Median“ statistisch signifikant und klinisch relevant aus (Tabelle 4-73). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Das Longitudinalplot für die Einzelkomponente Beweglichkeit der oberen Gliedmaßen zur Subgruppe „CRP zu Baseline (mg/l)“ ist in Abbildung 4-17 dargestellt.

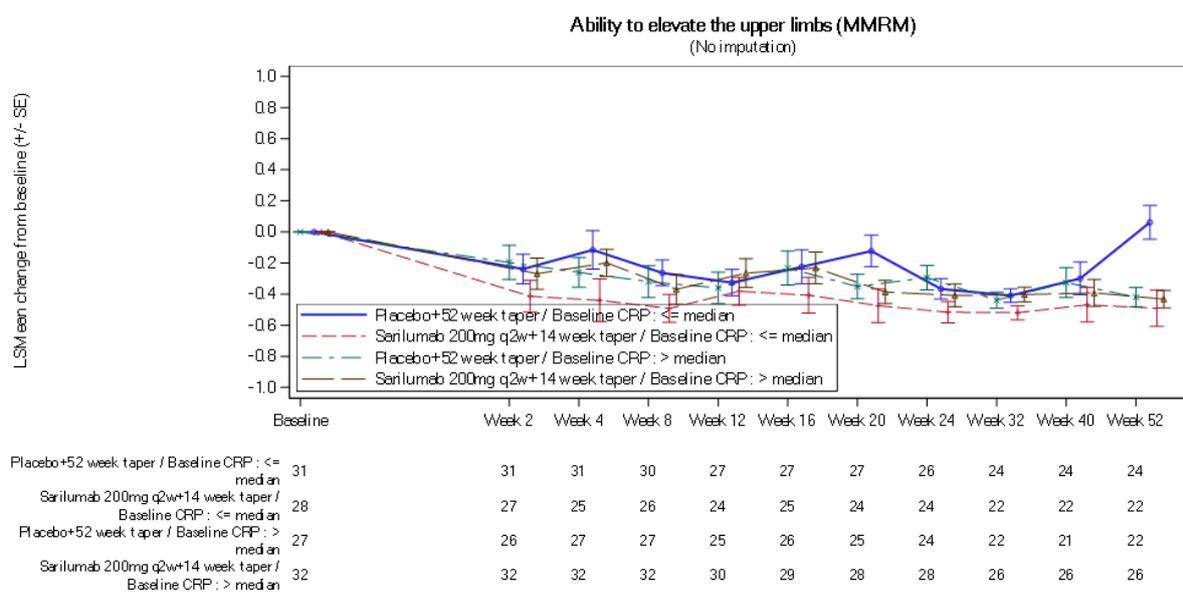


Abbildung 4-17: Longitudinalplot für die Veränderung der Einzelkomponente Beweglichkeit der oberen Gliedmaßen, Subgruppenmerkmal „CRP zu Baseline (mg/l)“, aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

CRP: C-reaktives Protein; ITT: Intention-To-Treat; LSMean: Kleinste-Quadrate-Mittelwert (Least Squares Mean); MMRM: Mixed Model for Repeated Measures; q2w: Alle zwei Wochen; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); SE: Standardfehler (Standard Error).

Quelle: Sanofi 2024b

Die nicht statistisch signifikanten Ergebnisse der Subgruppenanalysen der Veränderung des globalen PMR-AS und seiner Einzelkomponente CRP zu Woche 52 gegenüber Baseline mit den zugehörigen Longitudinalplots befinden sich ebenfalls in Anhang 4-G.

4.3.1.3.2.5 Subgruppenanalysen für den Endpunkt Veränderung des FACIT-F um $\geq 7,8$ Punkte – RCT

Tabelle 4-74: Ergebnisse der Subgruppenanalysen für die Veränderung des FACIT-F um $\geq 7,8$ Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions-test
Alle Patienten	24/60 (40,0)	17/58 (29,3)	1,61 [0,74; 3,49] 0,2267	1,36 [0,82; 2,28] 0,2331	10,69 [-6,55; 27,92] 0,2218	Nicht zutreffend
Altersgruppe (Jahre)						
< 70	18/34 (52,9)	6/25 (24,0)	3,56 [1,11; 11,40] 0,0328	2,21 [1,01; 4,83] 0,0476	28,94 [4,73; 53,16] 0,0200	0,0498
≥ 70	6/26 (23,1)	11/33 (33,3)	0,60 [0,18; 1,97] 0,3935	0,69 [0,29; 1,65] 0,4075	-10,26 [-33,58; 13,06] 0,3822	
CS-Dosis zu Baseline (mg)						
\leq Median	15/30 (50,0)	7/34 (20,6)	3,86 [1,26; 11,81] 0,0188	2,43 [1,13; 5,22] 0,0232	29,41 [6,50; 52,33] 0,0127	0,0241
> Median	9/30 (30,0)	10/24 (41,7)	0,60 [0,19; 1,90] 0,3781	0,72 [0,34; 1,51] 0,3845	-11,67 [-37,93; 14,59] 0,3768	
CS: Corticosteroids; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko.						
Quelle: Sanofi 2024d, 2024e						

Für den Endpunkt Veränderung des FACIT-F um $\geq 7,8$ Punkte ergaben sich statistisch signifikante Interaktionen zwischen der Behandlung und den subgruppenbildenden Merkmalen „Altersgruppe (Jahre)“ ($p=0,0498$) und „CS-Dosis zu Baseline (mg)“ ($p=0,0241$). Für das subgruppenbildende Merkmal „Altersgruppe (Jahre)“ zeigte sich in der Subgruppenausprägung „ ≥ 70 “ gegenüber der Gesamtpopulation ebenso ein Subgruppeneffekt wie für das subgruppenbildende Merkmal „CS-Dosis zu Baseline (mg)“ in der Subgruppenausprägung „ $>$ Median“. Es liegen qualitative Unterschiede vor. Die Effektschätzer in den Subgruppenausprägungen „ ≥ 70 “ und „ $>$ Median“ sind nicht gleichgerichtet zum Gesamteffekt. Das 95 %-KI des jeweiligen RR schließt jedoch die 1 mit ein. Für das subgruppenbildende Merkmal „Altersgruppe (Jahre)“ zeigte sich in der Subgruppenausprägung

„< 70“ gegenüber der Gesamtpopulation ebenso ein Subgruppeneffekt wie für das subgruppenbildende Merkmal „CS-Dosis zu Baseline (mg)“ in der Subgruppenausprägung „ \leq Median“. Es liegen statistisch signifikante Vorteile für Sarilumab vor (Tabelle 4-74). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Die Ergebnisse der Subgruppenanalysen der Veränderung des FACIT-F zu Woche 52 gegenüber Baseline mit den zugehörigen Longitudinalplots befinden sich in Anhang 4-G.

4.3.1.3.2.6 Subgruppenanalysen für den Endpunkt Veränderung der EQ-5D VAS – RCT

Für den Endpunkt EQ-5D VAS ergaben sich weder für die Veränderung um ≥ 15 Punkte noch für die Veränderung zu Woche 52 gegenüber Baseline zwischen der Behandlung und den subgruppenbildenden Merkmalen für eine Subgruppe statistisch signifikante Interaktionen. Die Ergebnisse der Subgruppenanalysen befinden sich in Anhang 4-G.

4.3.1.3.2.7 Subgruppenanalysen für den Endpunkt Veränderung des HAQ-DI – RCT

Tabelle 4-75: Ergebnisse der Subgruppenanalysen für die Veränderung des Standard-Scores des HAQ-DI um $\geq -0,45$ Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions-test
Alle Patienten	19/60 (31,7)	10/58 (17,2)	2,22 [0,92; 5,37] 0,0749	1,84 [0,93; 3,63] 0,0808	14,43 [-1,00; 29,85] 0,0666	Nicht zutreffend
Anzahl Rezidive vor Screening						
1	6/24 (25,0)	7/23 (30,4)	0,76 [0,20; 2,84] 0,6795	0,82 [0,32; 2,13] 0,6861	-5,43 [-31,71; 20,84] 0,6789	0,0337
> 1	13/36 (36,1)	3/35 (8,6)	6,03 [1,50; 24,19] 0,0120	4,21 [1,29; 13,80] 0,0175	27,54 [8,99; 46,09] 0,0042	
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); HAQ-DI: Health Assessment Questionnaire Disability Index; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko. Quelle: Sanofi 2024d, 2024e						

Für den Endpunkt Veränderung des Standard-Scores des HAQ-DI um $\geq -0,45$ Punkte ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „Anzahl der Rezidive vor Screening“ ($p=0,0337$). Für die Subgruppenausprägung „1“ zeigte sich gegenüber der Gesamtpopulation ein Subgruppeneffekt. Es liegt ein qualitativer Unterschied vor. Der Effektschätzer in der Subgruppenausprägung „1“ ist nicht gleichgerichtet zum Gesamteffekt. Das 95 %-KI des RR schließt jedoch die 1 mit ein. Für die Subgruppenausprägung „ > 1 “ zeigte sich gegenüber der Gesamtpopulation ebenfalls ein Subgruppeneffekt. Es liegt ein statistisch signifikanter Vorteil für Sarilumab vor (Tabelle 4-75). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Tabelle 4-76: Ergebnisse der Subgruppenanalysen für die Veränderung der globalen Krankheitsaktivität des HAQ-DI um ≥ -15 Punkte aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (ITT-Population)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktionstest
Alle Patienten	21/60 (35,0)	14/58 (24,1)	1,69 [0,75; 3,81] 0,2011	1,45 [0,81; 2,58] 0,2076	10,86 [-5,65; 27,37] 0,1951	Nicht zutreffend
Geografische Region						
Westliche Länder	18/47 (38,3)	7/44 (15,9)	3,28 [1,19; 9,03] 0,0220	2,41 [1,10; 5,26] 0,0274	22,39 [4,54; 40,24] 0,0145	0,0190
Rest der Welt	3/13 (23,1)	7/14 (50,0)	0,30 [0,05; 1,72] 0,1680	0,46 [0,14; 1,50] 0,1988	-26,92 [-63,48; 9,64] 0,1419	
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); HAQ-DI: Health Assessment Questionnaire Disability Index; ITT: Intention-To-Treat; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko. Quelle: Sanofi 2024d, 2024e						

Für den Endpunkt Veränderung der globalen Krankheitsaktivität des HAQ-DI um ≥ -15 Punkte ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „Geografische Region“ ($p=0,0190$). Für die Subgruppenausprägung „Rest der Welt“ zeigte sich gegenüber der Gesamtpopulation ein Subgruppeneffekt. Es liegt ein qualitativer Unterschied vor. Der Effektschätzer in der

Subgruppenausprägung „Rest der Welt“ ist nicht gleichgerichtet zum Gesamteffekt. Das 95 %-KI des RR schließt jedoch die 1 mit ein. Für die Subgruppenausprägung „Westliche Länder“ zeigte sich gegenüber der Gesamtpopulation ebenfalls ein Subgruppeneffekt. Es liegt ein statistisch signifikanter Vorteil für Sarilumab vor (Tabelle 4-76). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Die Ergebnisse der Subgruppenanalysen der Veränderung des HAQ-DI zu Woche 52 gegenüber Baseline mit den zugehörigen Longitudinalplots befinden sich ebenfalls in Anhang 4-G.

4.3.1.3.2.8 Subgruppenanalysen für den Endpunkt Veränderung der SF-36v2 – RCT

Für den Endpunkt SF-36v2 ergaben sich weder für die PCS noch die MCS für die Veränderung um ≥ 10 Punkte bzw. für die Veränderung zu Woche 52 gegenüber Baseline zwischen der Behandlung und den subgruppenbildenden Merkmalen für eine Subgruppe statistisch signifikante Interaktionen. Die Ergebnisse der Subgruppenanalysen befinden sich in Anhang 4-G.

4.3.1.3.2.9 Subgruppenanalysen für den Endpunkt Nebenwirkungen – RCT

Tabelle 4-77: Ergebnisse der Subgruppenanalysen für UE gesamt aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktionstest
Alle Patienten	56/59 (94,9)	49/58 (84,5)	3,43 [0,87; 13,58] 0,0788	1,12 [0,99; 1,27] 0,0711	10,43 [-0,56; 21,42] 0,0626	Nicht zutreffend
Geschlecht						
Männlich	14/14 (100,0)	17/21 (81,0)	NI	NI	NI	< 0,0001
Weiblich	42/45 (93,3)	32/37 (86,5)	2,19 [0,48; 10,07] 0,3106	1,08 [0,93; 1,26] 0,3249	6,85 [-6,56; 20,26] 0,3127	
Geografische Region						
Westliche Länder	46/46 (100,0)	41/44 (93,2)	NI	NI	NI	< 0,0001
Rest der Welt	10/13 (76,9)	8/14 (57,1)	2,50 [0,43; 14,44] 0,2921	1,35 [0,76; 2,38] 0,3069	19,78 [-16,57; 56,13] 0,2730	

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions-test
CRP zu Baseline (mg/l)						
≤ Median	27/27 (100,0)	26/31 (83,9)	NI	NI	NI	< 0,0001
> Median	29/32 (90,6)	23/27 (85,2)	1,68 [0,33; 8,57] 0,5255	1,06 [0,87; 1,30] 0,5379	5,44 [-11,70; 22,58] 0,5277	
ESR zu Baseline (mm/h)						
≤ Median	27/27 (100,0)	30/31 (96,8)	NI	NI	NI	< 0,0001
> Median	29/32 (90,6)	19/27 (70,4)	4,07 [0,93; 17,86] 0,0624	1,29 [0,98; 1,70] 0,0712	20,25 [-0,14; 40,65] 0,0516	
CRP: C-reaktives Protein; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ESR: Erythrozyten-Sedimentationsrate; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; NI: Nicht interpretierbar; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; UE: Unerwünschtes Ereignis. Quelle: Sanofi 2024h, 2024i						

Für den Endpunkt UE gesamt ergaben sich statistisch signifikante Interaktionen zwischen der Behandlung und den subgruppenbildenden Merkmalen „Geschlecht“ ($p < 0,0001$), „Geografische Region“ ($p < 0,0001$), „CRP zu Baseline (mg/l)“ ($p < 0,0001$) und „ESR zu Baseline (mm/h)“ ($p < 0,0001$). Da die Effektschätzer für jeweils eine Ausprägung der subgruppenbildenden Merkmale nicht interpretierbar sind (Tabelle 4-77), können potenzielle Effektunterschiede nicht beurteilt werden. Für die übrigen Subgruppen wurde keine statistisch signifikante Interaktion zwischen Behandlung und subgruppenbildenden Merkmalen festgestellt. Die Ergebnisse befinden sich in Anhang 4-G.

Für die SOC des Endpunkts UE gesamt mit statistisch signifikanten Unterschieden zwischen den Studienarmen, „Erkrankungen des Blutes und des Lymphsystems“ und „Erkrankungen des Gastrointestinaltrakts“ (Tabelle 4-59) traten für keine Subgruppe statistisch signifikante Interaktionen auf. Die Ergebnisse der Subgruppenanalysen befinden sich in Anhang 4-G.

Für den Endpunkt UE gesamt ergaben sich für keinen PT statistisch signifikante Unterschiede zwischen den Studienarmen (Tabelle 4-60).

Subgruppenanalysen entfallen für schwere UE, da für keine SOC und keinen PT Ereignisse bei mindestens 5 % der Patienten in einem Studienarm auftraten.

Tabelle 4-78: Ergebnisse der Subgruppenanalysen für SUE aus RCT (Studie SAPHYR) mit dem zu bewertenden Arzneimittel (Sicherheitspopulation)

Subgruppe	Sarilumab + CST	Placebo + CST	Sarilumab + CST versus Placebo + CST			
	n/N (%)	n/N (%)	OR [95 %-KI] p-Wert	RR [95 %-KI] p-Wert	RD [95 %-KI] p-Wert	p-Wert für Interaktions-test
Alle Patienten	8/59 (13,6)	12/58 (20,7)	0,60 [0,22; 1,62] 0,3108	0,66 [0,29; 1,50] 0,3164	-7,13 [-20,88; 6,62] 0,3063	Nicht zutreffend
ESR zu Baseline (mm/h)						
≤ Median	2/27 (7,4)	10/31 (32,3)	0,17 [0,03; 0,88] 0,0358	0,23 [0,05; 0,99] 0,0482	-24,85 [-44,47; -5,23] 0,0140	0,0258
> Median	6/32 (18,8)	2/27 (7,4)	2,88 [0,51; 16,25] 0,2248	2,53 [0,54; 11,91] 0,2399	11,34 [-5,77; 28,45] 0,1896	
CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); ESR: Erythrozyten-Sedimentationsrate; KI: Konfidenzintervall; N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; OR: Odds Ratio; RCT: Randomisierte, kontrollierte Studie (Randomized Controlled Trial); RD: Risikodifferenz; RR: Relatives Risiko; SUE: Schwerwiegendes unerwünschtes Ereignis. Quelle: Sanofi 2024h, 2024i						

Für den Endpunkt SUE ergab sich eine statistisch signifikante Interaktion zwischen der Behandlung und dem subgruppenbildenden Merkmal „ESR zu Baseline (mm/h)“ ($p=0,0258$). Für die Subgruppenausprägung „> Median“ zeigte sich gegenüber der Gesamtpopulation ein Subgruppeneffekt. Es liegt ein qualitativer Unterschied vor. Der Effektschätzer in der Subgruppenausprägung „> Median“ ist nicht gleichgerichtet zum Gesamteffekt. Das 95 %-KI des RR schließt jedoch die 1 mit ein. Für die Subgruppenausprägung „≤ Median“ zeigte sich gegenüber der Gesamtpopulation ebenfalls ein Subgruppeneffekt. Es liegt ein statistisch signifikanter Vorteil für Sarilumab vor (Tabelle 4-78). Aufgrund der Problematik des multiplen Testens kann eine scheinbare Effektmodifikation als Folge falsch positiver Interaktionstests nicht ausgeschlossen werden. Für die übrigen Merkmale kam es zu weniger als 10 Ereignissen in einer der Subgruppen.

Für den Endpunkt SUE wurden für keine SOC statistisch signifikante Unterschiede zwischen den Studienarmen beobachtet (Tabelle 4-61). PT traten bei weniger als 5 % der Patienten in einem Studienarm auf. Daher entfallen Subgruppenanalysen auf PT-Ebene.

Darüber hinaus erfolgen keine Subgruppenanalysen für UE, die zum Therapieabbruch führten, und UE von besonderem Interesse, da in beiden Studienarmen jeweils weniger als 10 Ereignisse auftraten.

4.3.1.4 Liste der eingeschlossenen Studien - RCT

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Tabelle 4-79: Übersicht der eingeschlossenen Studie SAPHYR

Studienbericht (ja/nein [Zitat])	Registereintrag (ja/nein [Zitat])	Publikation (ja/nein [Zitat])	Post-hoc-Analysen (ja/nein [Zitat])
ja (Sanofi 2021a)	ja Clinicaltrials.gov (ClinicalTrialsGOV 2022) EU-CTR (EU-CTR 0000) WHO-ICTRP (WHO- ICTRP 2021)	ja (Spiera 2023)	ja (Sanofi 2024b, 2024c, 2024d, 2024e, 2024f, 2024g, 2024h, 2024i)

4.3.2 Weitere Unterlagen

4.3.2.1 Indirekte Vergleiche auf Basis randomisierter kontrollierter Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn indirekte Vergleiche als Nachweis für einen Zusatznutzen herangezogen werden sollen. Das ist dann möglich, wenn keine direkten Vergleichsstudien für das zu bewertende Arzneimittel gegenüber der zweckmäßigen Vergleichstherapie vorliegen oder diese keine ausreichenden Aussagen über den Zusatznutzen zulassen.

4.3.2.1.1 Ergebnis der Informationsbeschaffung – Studien für indirekte Vergleiche

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu Studien für indirekte Vergleiche. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).** Benennen Sie sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken
- Studien aus der Suche auf der Internetseite des G-BA
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht zutreffend.

4.3.2.1.2 Charakteristika der Studien für indirekte Vergleiche

*Charakterisieren Sie nachfolgend die Studien, die für indirekte Vergleiche identifiziert wurden und bewerten Sie darüber hinaus deren Ähnlichkeit. Begründen Sie darauf basierend den Ein- bzw. Ausschluss von Studien für die von Ihnen durchgeführten indirekten Vergleiche. Bewerten Sie das Verzerrungspotenzial der für indirekte Vergleiche herangezogenen Studien. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

Nicht zutreffend.

4.3.2.1.3 Ergebnisse aus indirekten Vergleichen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus indirekten Vergleichen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-80: Matrix der Endpunkte in den eingeschlossenen RCT für indirekte Vergleiche

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend					

4.3.2.1.3.1 <Endpunkt xxx> – indirekte Vergleiche aus RCT

Für die indirekten Vergleiche soll zunächst für jeden Endpunkt eine Übersicht über die verfügbaren Vergleiche gegeben werden. Anschließend soll die Darstellung der Ergebnisse in drei Schritten erfolgen: 1) Bewertung des Verzerrungspotenzials auf Endpunktebene pro Studie, 2) tabellarische Darstellung der Ergebnisse der einzelnen Studien, 3) Darstellung des indirekten Vergleichs. **Für die Punkte 1 und 2 gelten die gleichen Anforderungen wie für die Darstellung der Ergebnisse der direkten Vergleiche in Abschnitt 0.**

Geben Sie für den im vorliegenden Abschnitt präsentierten Endpunkt einen Überblick über die in den Studien verfügbaren Vergleiche. Beispielfhaft wäre folgende Darstellung denkbar:

Tabelle 4-81: Zusammenfassung der verfügbaren Vergleiche in den Studien, die für den indirekten Vergleich herangezogen wurden

Anzahl Studien	Studie	Intervention	<Vergleichs-therapie 1>	<Vergleichs-therapie 2>	<Vergleichs-therapie 3>
Nicht zutreffend					

Stellen Sie zusätzlich die Netzwerkstruktur des indirekten Vergleichs grafisch dar.

Nicht zutreffend.

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-82: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
Nicht zutreffend	

Bewerten Sie das Verzerrungspotenzial für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-83: Bewertung des Verzerrungspotenzials für <Endpunkt xxx> in RCT für indirekte Vergleiche

Studie	Verzerrungspotenzial auf Studienebene	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte	Verzerrungspotenzial Endpunkt
Nicht zutreffend						

Begründen Sie für jede Studie die abschließende Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse für den Endpunkt xxx für jede einzelne Studie in tabellarischer Form dar. Fügen Sie für jede Studie eine neue Zeile ein. Beschreiben Sie die Ergebnisse zusammenfassend.

Tabelle 4-84: Ergebnisse für <Endpunkt xxx> aus RCT für indirekte Vergleiche

Studie	Tabellarische Präsentation in geeigneter Form (Anforderungen siehe Erläuterung in Abschnitt 0)
Nicht zutreffend	

Nicht zutreffend.

Stellen Sie die Ergebnisse der indirekten Vergleiche in tabellarischer Form dar. Optional können die Ergebnisse zusätzlich auch grafisch illustriert werden. Orientieren Sie sich dabei an der üblichen Darstellung metaanalytischer Ergebnisse. Gliedern Sie die Ergebnisse nach folgenden Punkten:

- *Homogenität der Ergebnisse: Stellen Sie die Ergebnisse der paarweisen Meta-Analysen dar. Diskutieren Sie das Ausmaß sowie die Gründe für das Auftreten der Heterogenität für alle direkten paarweisen Vergleiche.*
- *Ergebnisse zu den Effekten: Stellen Sie die gepoolten Ergebnisse dar.*
- *Konsistenzprüfung: Stellen Sie die Ergebnisse der Konsistenzprüfung dar. Diskutieren Sie insbesondere inkonsistente Ergebnisse.*

Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt für den ein indirekter Vergleich vorgenommen wird fortlaufend in einem eigenen Abschnitt dar.

4.3.2.1.3.2 Subgruppenanalysen – indirekte Vergleiche aus RCT

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen auf Basis indirekter Vergleiche aus RCT. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.**

Nicht zutreffend.

4.3.2.1.4 Liste der eingeschlossenen Studien – indirekte Vergleiche aus RCT

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.3.2.2 Nicht randomisierte vergleichende Studien

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn nicht randomisierte vergleichende Studien als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.2.1 Ergebnis der Informationsbeschaffung – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung zu nicht randomisierten vergleichenden Studien. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.).** Benennen Sie

- Studien des pharmazeutischen Unternehmers
- Studien aus der bibliografischen Literaturrecherche
- Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken
- Studien aus der Suche auf der G-BA Internetseite
- Resultierender Studienpool aus den einzelnen Suchschritten

Nicht zutreffend.

4.3.2.2.2 Charakteristika der nicht randomisierten vergleichenden Studien

*Charakterisieren Sie nachfolgend die nicht randomisierten vergleichenden Studien. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.***

Beschreiben Sie die Verzerrungsaspekte der nicht randomisierten vergleichenden Studie auf Studienebene mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Beschreibung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-85: Verzerrungsaspekte auf Studienebene – nicht randomisierte vergleichende Interventionsstudien

Studie	Zeitliche Parallelität der Gruppen	Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren	Verblindung		Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
			Patient	Behandelnde Personen		
Nicht zutreffend						

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Studienebene.

Nicht zutreffend.

4.3.2.2.3 Ergebnisse aus nicht randomisierten vergleichenden Studien

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus nicht randomisierten vergleichenden Studien beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-86: Matrix der Endpunkte in den eingeschlossenen nicht randomisierten vergleichenden Studien

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend					

4.3.2.2.3.1 <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-87: Operationalisierung von <Endpunkt xxx>

Studie	Operationalisierung
Nicht zutreffend	

Beschreiben Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt mithilfe des Bewertungsbogens in Anhang 4-F. Fassen Sie die Bewertung mit den Angaben in der folgenden Tabelle zusammen. Fügen Sie für jede Studie eine neue Zeile ein.

Dokumentieren Sie die Einschätzung für jede Studie mit einem Bewertungsbogen in Anhang 4-F.

Tabelle 4-88: Verzerrungsaspekte für <Endpunkt xxx> – nicht randomisierte vergleichende Studien

Studie	Verblindung Endpunkterheber	Adäquate Umsetzung des ITT-Prinzips	Ergebnisunabhängige Berichterstattung	Keine sonstigen Aspekte
Nicht zutreffend				

Beschreiben Sie zusammenfassend die Bewertungsergebnisse zu Verzerrungsaspekten auf Endpunktebene.

Nicht zutreffend.

Stellen Sie die Ergebnisse der nicht randomisierten vergleichenden Studien gemäß den Anforderungen des TREND- bzw. des STROBE-Statements dar. Machen Sie dabei auch Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus nicht randomisierten vergleichenden Studien fortlaufend in einem eigenen Abschnitt dar.

4.3.2.2.3.2 Subgruppenanalysen – nicht randomisierte vergleichende Studien

Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus nicht randomisierten vergleichenden Studien. Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.

Nicht zutreffend.

4.3.2.2.4 Liste der eingeschlossenen Studien – nicht randomisierte vergleichende Studien

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.3.2.3 Weitere Untersuchungen

Hinweis: Die nachfolgenden Unterabschnitte sind nur dann auszufüllen, wenn über die in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 genannten Studien hinausgehende Untersuchungen als Nachweis für einen Zusatznutzen herangezogen werden sollen.

4.3.2.3.1 Ergebnis der Informationsbeschaffung – weitere Untersuchungen

*Beschreiben Sie nachfolgend das Ergebnis der Informationsbeschaffung nach Untersuchungen, die nicht in den Abschnitten 4.3.1, 4.3.2.1 und 4.3.2.2 aufgeführt sind. **Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.1 (Ergebnis der Informationsbeschaffung – RCT mit dem zu bewertenden Arzneimittel) und stellen Sie Informationen sowohl für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie analog Abschnitt 4.3.1.1 zur Verfügung (einschließlich tabellarischer Darstellungen, Angabe eines Flussdiagramms etc.). Benennen Sie für das zu bewertende Arzneimittel als auch für die zweckmäßige Vergleichstherapie***

- *Studien des pharmazeutischen Unternehmers*
- *Studien aus der bibliografischen Literaturrecherche*
- *Studien aus der Suche in Studienregistern/ Studienergebnisdatenbanken*
- *Studien aus der Suche auf der G-BA Internetseite*
- *Resultierender Studienpool aus den einzelnen Suchschritten*

Nicht zutreffend.

4.3.2.3.2 Charakteristika der weiteren Untersuchungen

Charakterisieren Sie nachfolgend die weiteren Untersuchungen und bewerten Sie deren Verzerrungsaspekte.

Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Strukturieren Sie diesen Abschnitt analog Abschnitt 4.3.1.2 und stellen Sie Informationen analog Abschnitt 4.3.1.2 zur Verfügung.

Nicht zutreffend.

4.3.2.3.3 Ergebnisse aus weiteren Untersuchungen

Geben Sie in der folgenden Tabelle einen Überblick über die patientenrelevanten Endpunkte, auf denen Ihre Bewertung des medizinischen Nutzens und Zusatznutzens aus weiteren Untersuchungen beruht. Orientieren Sie sich dabei an der beispielhaften Angabe in der ersten Zeile. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-89: Matrix der Endpunkte in den eingeschlossenen weiteren Untersuchungen

Studie	<Mortalität>	<Gesundheits- bezogene Lebensqualität>	<Endpunkt>	<Endpunkt>	<Endpunkt>
Nicht zutreffend					

4.3.2.3.3.1 <Endpunkt xxx> – weitere Untersuchungen

Beschreiben Sie die Operationalisierung des Endpunkts für jede Studie in der folgenden Tabelle. Fügen Sie für jede Studie eine neue Zeile ein.

Tabelle 4-90: Operationalisierung von <Endpunkt xxx> – weitere Untersuchungen

Studie	Operationalisierung
Nicht zutreffend	

Bewerten Sie die Verzerrungsaspekte für den in diesem Abschnitt beschriebenen Endpunkt. Ergebnisse nicht randomisierter Studien, die keine kontrollierten Interventionsstudien sind, gelten aufgrund ihres Studiendesigns generell als potenziell hoch verzerrt. Trifft das auf die von Ihnen vorgelegten Studien nicht zu, begründen Sie Ihre Einschätzung.

Nicht zutreffend.

Stellen Sie die Ergebnisse der weiteren Untersuchungen gemäß den jeweils gültigen Standards für die Berichterstattung dar. Begründen Sie dabei die Auswahl des Standards für die Berichterstattung. Machen Sie darüber hinaus Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext.

Nicht zutreffend.

Stellen Sie die in diesem Abschnitt beschriebenen Informationen für jeden weiteren Endpunkt aus weiteren Untersuchungen fortlaufend in einem eigenen Abschnitt dar.

4.3.2.3.3.2 Subgruppenanalysen – weitere Untersuchungen

*Beschreiben Sie nachfolgend die Ergebnisse von Subgruppenanalysen aus weiteren Untersuchungen. **Berücksichtigen Sie dabei die Anforderungen gemäß Abschnitt 4.3.1.3.2.***

Nicht zutreffend.

4.3.2.3.4 Liste der eingeschlossenen Studien – weitere Untersuchungen

Listen Sie alle für die Nutzenbewertung berücksichtigten Studien und Untersuchungen unter Angabe der im Dossier verwendeten Studienbezeichnung und der zugehörigen Quellen (z. B. Publikationen, Studienberichte, Studienregistereinträge).

Nicht zutreffend.

4.4 Abschließende Bewertung der Unterlagen zum Nachweis des Zusatznutzens

4.4.1 Beurteilung der Aussagekraft der Nachweise

Legen Sie für alle im Dossier eingereichten Unterlagen die Evidenzstufe dar. Beschreiben Sie zusammenfassend auf Basis der in den Abschnitten 4.3.1 und 4.3.2 präsentierten Ergebnisse die Aussagekraft der Nachweise für einen Zusatznutzen unter Berücksichtigung der Studienqualität, der Validität der herangezogenen Endpunkte sowie der Evidenzstufe.

Sarilumab ist gemäß Fachinformation indiziert für die Behandlung der Polymyalgia rheumatica (PMR) bei erwachsenen Patienten, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt (Sanofi 2024a).

Als zVT von Sarilumab im vorliegenden AWG bestimmte der G-BA eine „Therapie nach ärztlicher Maßgabe unter Berücksichtigung von systemischen Glukokortikoiden und der Kombination von Glukokortikoiden mit Methotrexat“ (G-BA 2024).

Zum Nachweis des Nutzens und Zusatznutzens von Sarilumab liegt die klinische Zulassungsstudie SAPHYR vor. Bei der Studie handelt es sich um eine abgeschlossene, randomisierte, multizentrische, doppelblinde, placebokontrollierte, parallele Phase-III-Studie an Patienten im Alter ab 50 Jahren mit aktiver PMR, die mit Prednison vorbehandelt waren und zu Studienbeginn eine Prednison-Dosis von mindestens 7,5 mg/Tag und höchstens 20 mg/Tag oder Äquivalent einnahmen. Die Patienten wurden im Verhältnis 1:1 entweder Sarilumab 200 mg q2w oder Placebo zu Sarilumab q2w zugeteilt. Zusätzlich erhielten die Patienten Prednison, das im Behandlungsverlauf über 14 Wochen im Sarilumab-Arm bzw. über

52 Wochen im Kontrollarm ausgeschlichen wurde. Zur Aufrechterhaltung der Verblindung wurde den Patienten Prednison, Placebo zu Prednison oder eine Kombination aus Prednison und Placebo zu Prednison verabreicht. Anstelle der verblindeten Gabe von Prednison bzw. Placebo zu Prednison konnten kommerzielle Corticosteroide, einschließlich Prednison, in beiden Studienarmen von Woche 12 bis Woche 51 als Notfallmedikation verabreicht werden.

Die mediane Behandlungsdauer lag in beiden Studienarmen bei 363,0 Tagen. Ein Patient, der in den Sarilumab-Arm randomisiert wurde, erhielt keine Studienmedikation. Von den übrigen 59 Patienten beendeten 42 (70,0 %) die Behandlung protokollgerecht, während 17 Patienten (28,3 %) sie frühzeitig beendeten. Im Kontrollarm beendeten 36 Patienten (62,1 %) die Behandlung protokollgerecht, 22 (37,9 %) brachen sie frühzeitig ab. Als häufigster Grund für den Behandlungsabbruch wurden im Sarilumab-Arm unerwünschte Ereignisse bei 11,7 % der Patienten, im Kontrollarm mangelnde Wirksamkeit bei 15,5 % der Patienten angegeben (Sanofi 2021a).

Die Studie SAPHYR entspricht nach der in der Verfahrensordnung des G-BA dargestellten Evidenzhierarchie der Evidenzstufe Ib (G-BA 2025). Das Verzerrungspotenzial auf Studienebene ist insgesamt als niedrig einzustufen. Randomisierung und Gruppenzuteilung erfolgten adäquat und verdeckt mit Hilfe einer IRT. Sowohl Patienten als auch behandelnde Personen waren hinsichtlich der Zuteilung zu Sarilumab oder Placebo verblindet. Eine ergebnisabhängige Berichterstattung kann ausgeschlossen werden, da die Analysen prospektiv im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden (G-BA 2025). Darüber hinaus lagen keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Es kann daher von einer hohen Studienqualität der Studie SAPHYR ausgegangen werden.

Der Nachweis des Zusatznutzens von Sarilumab im vorliegenden AWG wird anhand der Studie SAPHYR über einen direkten Vergleich von Sarilumab + CST gegenüber Placebo + CST geführt. Die Studienpopulation ist eine Teilpopulation des Labels, d. h., es handelt sich um die Patienten, bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt, und wird im Dossier zur Gewährleistung einer einheitlichen zVT nicht weiter unterteilt dargestellt. Es lag in der Entscheidung der Prüfarzte, den Patienten bei Bedarf in den ersten 12 Wochen der Behandlung zusätzliches Prednison zu verabreichen. Darüber hinaus hatten sie die Möglichkeit, diejenigen Patienten, die während der letzten 3 Monate vor Studienbeginn auf eine stabile Dosis von MTX eingestellt waren, im gesamten Studienverlauf mit MTX weiter zu behandeln. MTX konnte zudem als Notfallmedikation eingesetzt werden, sofern Corticosteroide keine ausreichende Wirksamkeit zeigten. Dieses Vorgehen entspricht den Empfehlungen der S3-Leitlinie (Buttgereit 2018). Insgesamt nahmen 29 Patienten (24,6 %), davon 17 (29,3 %) im Kontrollarm, begleitend MTX ein. Es bestand keine Effektmodifikation durch die MTX-Gabe, da sich für das Subgruppenmerkmal „MTX-Add-on-Therapie“ mit den Ausprägungen „Ja“ und „Nein“ für keinen Endpunkt statistisch signifikante Interaktionen ergaben (siehe Tabelle 4-65). Die zVT wird in der Studie SAPHYR somit als adäquat umgesetzt betrachtet.

Für die Aussagen zum Zusatznutzen von Sarilumab für die Behandlung von erwachsenen Patienten mit aktiver PMR wird gegenüber der zVT, basierend auf dem im Dossier vorgelegten

direkten Vergleich mit Placebo + CST, bezüglich Aussagekraft und Ergebnissicherheit der Status eines Hinweises beansprucht.

4.4.2 Beschreibung des Zusatznutzens einschließlich dessen Wahrscheinlichkeit und Ausmaß

Führen Sie die in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse zum Zusatznutzen auf Ebene einzelner Endpunkte zusammen und leiten Sie ab, ob sich aus der Zusammenschau der Ergebnisse zu den einzelnen Endpunkten insgesamt ein Zusatznutzen des zu bewertenden Arzneimittels im Vergleich zur zweckmäßigen Vergleichstherapie ergibt. Berücksichtigen Sie dabei auch die Angaben zur Übertragbarkeit der Studienergebnisse auf den deutschen Versorgungskontext. Liegt ein Zusatznutzen vor, beschreiben Sie worin der Zusatznutzen besteht.

Stellen Sie die Wahrscheinlichkeit des Zusatznutzens dar, d. h., beschreiben und begründen Sie unter Berücksichtigung der in Abschnitt 4.4.1 dargelegten Aussagekraft der Nachweise die Ergebnissicherheit der Aussage zum Zusatznutzen.

Beschreiben Sie außerdem das Ausmaß des Zusatznutzens unter Verwendung folgender Kategorisierung (in der Definition gemäß AM-NutzenV):

- *erheblicher Zusatznutzen*
- *beträchtlicher Zusatznutzen*
- *geringer Zusatznutzen*
- *nicht quantifizierbarer Zusatznutzen*
- *kein Zusatznutzen belegbar*
- *der Nutzen des zu bewertenden Arzneimittels ist geringer als der Nutzen der zweckmäßigen Vergleichstherapie*

Berücksichtigen Sie bei den Aussagen zum Zusatznutzen ggf. nachgewiesene Unterschiede zwischen verschiedenen Patientengruppen.

Die Bewertung des medizinischen Zusatznutzens erfolgt unter Anwendung der Kriterien der AM-NutzenV für die Nutzendimensionen Mortalität, Morbidität, gesundheitsbezogene Lebensqualität und Nebenwirkungen (BMJ 2023).

Tabelle 4-91: Ergebnisse, Aussagesicherheit und Ausmaß des Zusatznutzens von Sarilumab gegenüber der zVT auf Endpunktebene

Nutzendimension/ Endpunkt	Sarilumab+CST n/N (%)	Placebo+CST n/N (%)	Sarilumab+CST vs. Placebo+CST	Aussagesicherheit und Ausmaß des Zusatznutzens
			Effektschätzer [95 %-KI]; p-Wert	
Gesamtmortalität				
Keine Todesfälle in der Studie SAPHYR				Nicht zur Ableitung des Zusatznutzens herangezogen
Morbidität				
Remission				
Anhaltende Remission zu Woche 52	17/60 (28,3)	6/58 (10,3)	RR 2,74 [1,15; 6,52]; 0,0227	Hinweis auf einen beträchtlichen Zusatznutzen
Einzelkomponenten der Remission				
Remission spätestens zu Woche 12	28/60 (46,7)	22/58 (37,9)	RR 1,23 [0,80; 1,89]; 0,3454	Zusatznutzen nicht belegt
Kein Krankheitsschub von Woche 12 bis 52	33/60 (55,0)	19/58 (32,8)	RR 1,68 [1,08; 2,60]; 0,0206	Hinweis auf einen geringen Zusatznutzen
Anhaltende CRP- Reduktion von Woche 12 bis 52	40/60 (66,7)	26/58 (44,8)	RR 1,49 [1,06; 2,09]; 0,0223	Nicht zur Ableitung des Zusatznutzens herangezogen
Erfolgreiches Prednison- Ausschleichen von Woche 12 bis 52	30/60 (50,0)	14/58 (24,1)	RR 2,07 [1,22; 3,51]; 0,0068	Hinweis auf einen beträchtlichen Zusatznutzen
PMR-Schub				
Zeit bis zum ersten PMR- Schub nach klinischer Remission	29/60 (48,3)	45/58 (77,6)	HR 0,56 [0,351; 0,895]; 0,0153	Hinweis auf einen beträchtlichen Zusatznutzen
PMR-AS				
Veränderung des PMR- AS zu Woche 52	31/60 (51,7)	28/58 (48,3)	LS-MWD -5,30 [-8,01; -2,60]; 0,0002 Hedges' g -1,01 [-1,52; -0,49]	Anhaltspunkt für einen beträchtlichen Zusatznutzen
Einzelkomponenten des PMR-AS				
Veränderung der VAS – Schmerz, Patientenbeurteilung (0 bis 10) um $\geq -1,5$ Punkte	20/60 (33,3)	20/58 (34,5)	RR 0,97 [0,58; 1,61]; 0,8962	Zusatznutzen nicht belegt
Veränderung der VAS – Krankheitsaktivität,	30/60 (50,0)	21/58 (36,2)	RR 1,38 [0,90; 2,12]; 0,1408	Zusatznutzen nicht belegt

Nutzendimension/ Endpunkt	Sarilumab+CST n/N (%)	Placebo+CST n/N (%)	Sarilumab+CST vs. Placebo+CST	Aussagesicherheit und Ausmaß des Zusatznutzens
			Effektschätzer [95 %-KI]; p-Wert	
Prüfarztbeurteilung (0 bis 10) um $\geq -1,5$ Punkte				
Veränderung des CRP (mg/dl) zu Woche 52	43/60 (71,7)	35/58 (60,3)	LS-MWD -7,10 [-11,91; -2,29]; 0,0042 Hedges' g -0,66 [-1,10; -0,21]	Nicht zur Ableitung des Zusatznutzens herangezogen
Veränderung der Dauer der Morgensteifigkeit (min x 0,1) zu Woche 52	48/60 (80,0)	46/58 (79,3)	LS-MWD -22,43 [-39,09; -5,77]; 0,0088 Hedges' g -0,55 [-0,96; -0,14]	Zusatznutzen nicht belegt
Veränderung der Beweglichkeit der oberen Gliedermaßen (EUL [3-0]) zu Woche 52	48/60 (80,0)	46/58 (79,3)	LS-MWD -0,24 [-0,40; -0,08]; 0,0044 Hedges' g -0,60 [-1,00; -0,19]	Zusatznutzen nicht belegt
Gesundheitszustand				
Veränderung des FACIT- F um $\geq 7,8$ Punkte	24/60 (40,0)	17/58 (29,3)	RR 1,36 [0,82; 2,28]; 0,2331	Zusatznutzen nicht belegt
Veränderung der EQ- 5D VAS um ≥ 15 Punkte	16/60 (26,7)	9/58 (15,5)	RR 1,72 [0,82; 3,60]; 0,1517	Zusatznutzen nicht belegt
Körperlicher Funktionsstatus				
Veränderung des HAQ- DI um $\geq -0,45$ Punkte	19/60 (31,7)	10/58 (17,2)	RR 1,84 [0,93; 3,63]; 0,0808	Zusatznutzen nicht belegt
Gesundheitsbezogene Lebensqualität				
Veränderung des SF-36v2				
Veränderung der PCS um ≥ 10 Punkte	14/60 (23,3)	10/58 (17,2)	RR 1,35 [0,65; 2,82]; 0,4194	Zusatznutzen nicht belegt
Veränderung der MCS um ≥ 10 Punkte	11/60 (18,3)	5/58 (8,6)	RR 2,13 [0,78; 5,80]; 0,1408	Zusatznutzen nicht belegt
Nebenwirkungen				
UE gesamt	56/59 (94,9)	49/58 (84,5)	RR 1,1 [0,99; 1,27]; 0,0711	Nicht zur Ableitung des Zusatznutzens herangezogen
SUE	8/59 (13,6)	12/58 (20,7)	RR 0,66 [0,29; 1,50]; 0,3164	Zusatznutzen nicht belegt

Medizinischer Nutzen, medizinischer Zusatznutzen, Patientengruppen mit therap. bedeutsamem Zusatznutzen

Nutzendimension/ Endpunkt	Sarilumab+CST n/N (%)	Placebo+CST n/N (%)	Sarilumab+CST vs. Placebo+CST	Aussagesicherheit und Ausmaß des Zusatznutzens
			Effektschätzer [95 %-KI]; p-Wert	
Schwere UE	5/59 (8,5)	7/58 (12,1)	RR 0,70 [0,23; 2,11]; 0,5289	Zusatznutzen nicht belegt
UE, die zum Therapieabbruch führten	7/59 (11,9)	4/58 (6,9)	RR 1,72 [0,53; 5,63]; 0,3701	Zusatznutzen nicht belegt
UE von besonderem Interesse	7/59 (11,9)	5/58 (8,6)	RR 1,38 [0,46; 4,14]; 0,5695	Zusatznutzen nicht belegt

CRP: C-reaktives Protein; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); EQ-5D: European Quality of Life 5 Dimensions; EULAR: European League Against Rheumatism (EULAR); FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI: Health Assessment Questionnaire Disability Index; HR: Hazard Ratio; KI: Konfidenzintervall; LS-MWD: Kleinste-Quadrat-Mittelwertdifferenz (Least Square-Mean Value Difference); MCS: Summenskala der psychischen Gesundheit (Mental Health Component Summary Score); N: Anzahl der Patienten; n: Anzahl der Patienten mit Ereignis; PCS: Summenskala der körperlichen Gesundheit (Physical Health Component Summary Score); PMR: Polymyalgia rheumatica; PMR-AS: PMR-Aktivitätsscore; RR: Relatives Risiko; SF-36v2: 36-Item Short Form Health Survey, Version 2; SUE: Schwerwiegendes UE; UE: Unerwünschtes Ereignis; VAS: Visuelle Analogskala; zVT: Zweckmäßige Vergleichstherapie.

Tabelle 4-92: Ergebnis, Aussagesicherheit und Ausmaß des Zusatznutzens von Sarilumab gegenüber der zVT für den Endpunkt Kumulative CS-Gesamtdosis

Nutzendimension/ Endpunkt	Sarilumab+CST Median (mg) [95 %-KI]	Placebo+CST Median (mg) [95 %-KI]	Sarilumab+CST vs. Placebo+CST	Aussagesicherheit und Ausmaß des Zusatznutzens
			p-Wert	
Morbidität				
Kumulative CS-Gesamtdosis				
Differenz zwischen tatsächlicher und erwarteter Gesamtdosis	0,0 [0,0; 0,0]	199,5 [0,0; 496,0]	0,0189	Nicht zur Ableitung des Zusatznutzens herangezogen

CS: Corticosteroids; CST: Corticosteroid-Ausschleichen (Corticosteroid Taper); KI: Konfidenzintervall;
zVT: Zweckmäßige Vergleichstherapie.

Die Ableitung des Zusatznutzens erfolgt für alle Patienten auf Basis der Gesamtpopulation (Hauptanalyse). Subgruppeneffekte werden gemäß den Vorgaben der Verfahrensordnung des G-BA in Abschnitt 4.3.1.3.2 dargestellt. Es konnten jedoch keine endpunktübergreifenden Effektmodifikationen festgestellt werden. Bei den vereinzelt beobachteten Subgruppeneffekten kann nicht abschließend bestimmt werden, ob es sich um eine tatsächliche Effektmodifikation oder um ein falsch-positives Ergebnis des Interaktionstests handelt. Somit bleiben die Subgruppenanalysen zur Ableitung des Zusatznutzens unberücksichtigt.

Mortalität

Die Kategorie Mortalität kann zur Ableitung des Zusatznutzens nicht herangezogen werden, da in der Studie SAPHYR keine Todesfälle auftraten.

Morbidität

Anhaltende Remission zu Woche 52

Für den Gesamtpunkt „anhaltende Remission zu Woche 52“ zeigte sich bezüglich des RR der Responderanalyse ein statistisch signifikanter und klinisch relevanter Vorteil von Sarilumab + CST im Vergleich zur zVT (RR 2,74; 95 %-KI [1,15; 6,52]). Dieses Ergebnis wird insbesondere durch die Einzelkomponente „erfolgreiches Prednison-Ausschleichen von Woche 12 bis 52“ gestützt (RR 2,07; 95 %-KI [1,22; 3,51]). Auch die beiden Einzelkomponenten „kein Krankheitsschub von Woche 12 bis 52“ und „anhaltende CRP-Reduktion von Woche 12 bis 52“ zeigen statistisch signifikante und klinisch relevante Ergebnisse zugunsten von Sarilumab + CST (RR 1,68; 95 %-KI [1,08; 2,60] bzw. RR 1,49; 95 %-KI [1,06; 2,09]). Die anhaltende CRP-Reduktion wird nicht zur Ableitung des Zusatznutzens herangezogen, da CRP als Laborparameter keine Patientenrelevanz besitzt. Für die Einzelkomponente „Remission spätestens zu Woche 12“ zeigte sich ein numerischer, aber statistisch nicht signifikanter, Vorteil für Sarilumab + CST gegenüber der zVT. Aus den Ergebnissen ergibt sich für den Gesamtpunkt „anhaltende Remission zu Woche 52“ gemäß § 5 Absatz 7 der AM-NutzenV ein Hinweis auf einen beträchtlichen Zusatznutzen.

PMR-Schub

Für den patientenrelevanten Endpunkt „Zeit bis zum ersten PMR-Schub nach klinischer Remission“ liegt bezüglich des HR eine statistisch signifikante und klinisch relevante Überlegenheit von Sarilumab + CST gegenüber der zVT basierend auf dem Cox-Regressionsmodell vor (HR 0,56; 95 %-KI [0,351; 0,895]). Ein PMR-Rezidiv wurde unter Sarilumab von deutlich weniger Patienten und später verzeichnet als unter der zVT. Je länger sich ein PMR-Schub hinauszögern lässt, desto später müssen Corticosteroide wieder bzw. in niedriger Prä-Rezidivdosis gegeben werden. Die auf diese Weise erreichte Verringerung der kumulativen Gesamtdosis bzw. Vermeidung der steroid-bedingten Nebenwirkungen bedeutet für die Patienten eine spürbare Verbesserung der gesundheitsbezogenen Lebensqualität. Das Ergebnis entspricht gemäß § 5 Absatz 7 der AM-NutzenV einem Hinweis auf einen beträchtlichen Zusatznutzen.

PMR-AS

Für den Gesamtpunkt „Veränderung des PMR-AS zu Woche 52“ ergab sich basierend auf dem MMRM-Modell eine statistisch signifikante und klinisch relevante Überlegenheit von Sarilumab + CST gegenüber der zVT (Hedges' g -1,01; 95 %-KI [-1,52; -0,49]). Einen ebenfalls statistisch signifikanten und klinisch relevanten Vorteil für Sarilumab + CST gegenüber der zVT zeigte die Einzelkomponente „Veränderung des CRP zu Woche 52“, die jedoch zur Ableitung des Zusatznutzens nicht berücksichtigt wird. Statistisch signifikante Unterschiede zugunsten von Sarilumab + CST wurden auch für die Einzelkomponenten „Veränderung der Dauer der Morgensteifigkeit zu Woche 52“ und „Veränderung der

Beweglichkeit der oberen Gliedmaßen zu Woche 52“ beobachtet, klinische Relevanz konnte jedoch nicht belegt werden. Für die beiden übrigen Einzelkomponenten „Veränderung der VAS - Schmerz, Patientenbeurteilung (0 bis 10) um $\geq -1,5$ Punkte“ sowie „Veränderung der VAS - Krankheitsaktivität, Prüfarztbeurteilung (0 bis 10) um $\geq -1,5$ Punkte“ fielen die Unterschiede zwischen Sarilumab + CST und der zVT nicht statistisch signifikant aus, die Prüfarztbeurteilung der Krankheitsaktivität ließ jedoch einen numerischen Vorteil für Sarilumab + CST erkennen. Insgesamt ergibt sich aus den Ergebnissen für den Endpunkt „Veränderung des PMR-AS zu Woche 52“ gemäß § 5 Absatz 7 der AM-NutzenV ein Anhaltspunkt für einen beträchtlichen Zusatznutzen.

Gesundheitszustand

Im Sarilumab-Arm erreichten mehr Patienten eine Verbesserung des Gesundheitszustands gemäß den Response-Kriterien des FACIT-F und der EQ-5D VAS, der Veränderung um $\geq 7,8$ Punkte bzw. ≥ 15 Punkte, als im Kontrollarm. Es zeigte sich ein numerischer Vorteil für Sarilumab + CST gegenüber der zVT. Die Unterschiede fielen jedoch nicht statistisch signifikant aus. Für die beiden Endpunkte zum Gesundheitsstatus „Veränderung des FACIT-F um $\geq 7,8$ Punkte“ und „Veränderung der EQ-5D VAS um ≥ 15 Punkte“ kann kein Zusatznutzen abgeleitet werden.

Körperlicher Funktionsstatus

Hinsichtlich des körperlichen Funktionsstatus gemäß HAQ-DI verbesserten sich im Sarilumab-Arm fast doppelt so viele Patienten entsprechend dem Response-Kriterium, der Veränderung um $\geq -0,45$ Punkte, wie im Kontrollarm. Daraus ergab sich ein numerischer Vorteil für Sarilumab + CST gegenüber der zVT, ohne dass statistische Signifikanz vorlag. Für den Endpunkt zum körperlichen Funktionsstatus „Veränderung des HAQ-DI um $\geq -0,45$ Punkte“ kann kein Zusatznutzen abgeleitet werden.

Kumulative CS-Gesamtdosis

Für den Endpunkt „Kumulative CS-Gesamtdosis“ ergab sich für Sarilumab + CST mit einer medianen Differenz zwischen tatsächlicher und erwarteter Gesamtdosis von 0,0 mg ein statistisch signifikanter Vorteil gegenüber der zVT mit einer medianen Differenz von 199,5 mg. Dies verdeutlicht, dass Sarilumab den Bedarf sowie die zusätzliche Gabe von CS signifikant verringert und somit steroidbedingte Nebenwirkungen entscheidend reduzieren kann. Aus dem Ergebnis kann kein Zusatznutzen abgeleitet werden.

Gesundheitsbezogene Lebensqualität

Für die gesundheitsbezogene Lebensqualität gemäß SF-36v2 zeigte das Response-Kriterium, die Veränderung um ≥ 10 Punkte, sowohl für die PCS als auch die MCS einen numerischen, aber nicht statistisch signifikanten, Vorteil für Sarilumab + CST gegenüber der zVT. Für die Endpunkte zur gesundheitsbezogenen Lebensqualität gemäß SF-36v2 „Veränderung der PCS um ≥ 10 Punkte“ und „Veränderung der MCS um ≥ 10 Punkte“ kann kein Zusatznutzen abgeleitet werden.

Nebenwirkungen

Die Ergebnisse der Sicherheitsendpunkte fielen für die beiden Studienarme insgesamt vergleichbar aus. Für UE gesamt, UE, die zum Therapieabbruch führten, und UE von besonderem Interesse wurden numerische Nachteile und für SUE und schwere UE numerische Vorteile für Sarilumab + CST gegenüber der zVT verzeichnet. Die Unterschiede waren nicht statistisch signifikant. Die Betrachtung der SOC und PT, die nicht gesondert zur Ableitung des Zusatznutzens herangezogen werden, ergab ein insgesamt ausgewogenes Bild für die beiden Studienarme. Statistisch signifikante Unterschiede zeigten sich lediglich für 3 SOC bei den UE gesamt, die für „Erkrankungen des Blutes und des Lymphsystems“ und „Erkrankungen des Gastrointestinaltrakts“ zuungunsten von Sarilumab + CST ausfielen. Für den Endpunkt „Nebenwirkungen“ ergibt sich weder ein Zusatznutzen noch ein geringerer Nutzen für Sarilumab + CST.

Gesamtschau

Zusammenfassend zeigen sich statistisch signifikante Vorteile beim primären Studienendpunkt „anhaltende Remission zu Woche 52“ sowie den sekundären Endpunkten „Zeit bis zum ersten PMR-Schub nach klinischer Remission“ und „Veränderung des PMR-AS zu Woche 52“. In der Gesamtschau ergibt sich demnach unter Berücksichtigung dieser Vorteile ein Hinweis auf einen beträchtlichen Zusatznutzen für Sarilumab + CST gegenüber der zVT in der Behandlung erwachsener Patienten mit PMR, die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt.

4.4.3 Angabe der Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht

Geben Sie auf Basis der in den Abschnitten 4.3.1 und 4.3.2 beschriebenen Ergebnisse und unter Berücksichtigung des in Abschnitt 4.4.2 dargelegten Zusatznutzens sowie dessen Wahrscheinlichkeit und Ausmaß in der nachfolgenden Tabelle an, für welche Patientengruppen ein therapeutisch bedeutsamer Zusatznutzen besteht. Benennen Sie das Ausmaß des Zusatznutzens in Patientengruppen mit therapeutisch bedeutsamem Zusatznutzen. Fügen Sie für jede Patientengruppe mit therapeutisch bedeutsamem Zusatznutzen eine neue Zeile ein.

Tabelle 4-93: Patientengruppen, für die ein therapeutisch bedeutsamer Zusatznutzen besteht, einschließlich Ausmaß des Zusatznutzens

Bezeichnung der Patientengruppen	Ausmaß des Zusatznutzens
Erwachsene Patienten mit Polymyalgia rheumatica (PMR), die auf Corticosteroide unzureichend angesprochen haben oder bei denen ein Rezidiv während des Ausschleichens der Corticosteroide auftritt	Hinweis auf einen beträchtlichen Zusatznutzen

4.5 Begründung für die Vorlage weiterer Unterlagen und Surrogatendpunkte

4.5.1 Begründung für die Vorlage indirekter Vergleiche

Sofern mit dem Dossier indirekte Vergleiche (Abschnitt 4.3.2.1) eingereicht wurden, begründen Sie dies. Begründen Sie dabei auch, warum sich die ausgewählten Studien jeweils für einen indirekten Vergleich gegenüber dem zu bewertenden Arzneimittel und damit für den Nachweis eines Zusatznutzens durch indirekten Vergleich eignen.

Nicht zutreffend.

4.5.2 Begründung für die Vorlage nicht randomisierter vergleichender Studien und weiterer Untersuchungen

Sofern mit dem Dossier nicht randomisierte vergleichende Studien (Abschnitt 4.3.2.2) oder weitere Untersuchungen (Abschnitt 4.3.2.3) eingereicht wurden, nennen Sie die Gründe, nach denen es unmöglich oder unangemessen ist, zu den in diesen Studien bzw. Untersuchungen behandelten Fragestellungen Studien höchster Evidenzstufe (randomisierte klinische Studien) durchzuführen oder zu fordern.

Nicht zutreffend.

4.5.3 Begründung für die Bewertung auf Grundlage der verfügbaren Evidenz, da valide Daten zu patientenrelevanten Endpunkten noch nicht vorliegen

Falls aus Ihrer Sicht valide Daten zu patientenrelevanten Endpunkten zum Zeitpunkt der Bewertung noch nicht vorliegen können, begründen Sie dies.

Nicht zutreffend.

4.5.4 Verwendung von Surrogatendpunkten

Die Verwendung von Surrogatendpunkten bedarf einer Begründung (siehe Abschnitt 4.5.3). Zusätzlich soll dargelegt werden, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Eine Validierung von Surrogatendpunkten bedarf in der Regel einer Meta-Analyse von Studien, in denen sowohl Effekte auf den Surrogatendpunkt als auch Effekte auf den interessierenden

patientenrelevanten Endpunkt untersucht wurden (Burzykowski 2005¹⁷, Molenberghs 2010¹⁸). Diese Studien müssen bei Patientenkollektiven und Interventionen durchgeführt worden sein, die Aussagen für das dem vorliegenden Antrag zugrundeliegende Anwendungsgebiet und das zu bewertende Arzneimittel sowie die Vergleichstherapie erlauben.

Eine Möglichkeit der Verwendung von Surrogatendpunkten ohne abschließende Validierung stellt die Anwendung des Konzepts eines sogenannten Surrogate-Threshold-Effekts (STE) (Burzykowski 2006¹⁹) dar. Daneben besteht die Möglichkeit einer Surrogatvalidierung in der quantitativen Betrachtung geeigneter Korrelationsmaße von Surrogatendpunkt und interessierendem patientenrelevanten Endpunkt („individuelle Ebene“) sowie von Effekten auf den Surrogatendpunkt und Effekten auf den interessierenden patientenrelevanten Endpunkt („Studienebene“). Dabei ist dann zu zeigen, dass die unteren Grenzen der entsprechenden 95%- Konfidenzintervalle für solche Korrelationsmaße ausreichend hoch sind. Die Anwendung alternativer Methoden zur Surrogatvalidierung (siehe Weir 2006²⁰) soll ausreichend begründet werden, insbesondere dann, wenn als Datengrundlage nur eine einzige Studie verwendet werden soll.

Berichten Sie zu den Studien zur Validierung oder zur Begründung für die Verwendung von Surrogatendpunkten mindestens folgende Informationen:

- Patientenpopulation
- Intervention
- Kontrolle
- Datenherkunft
- verwendete Methodik
- entsprechende Ergebnisse (abhängig von der Methode)
- Untersuchungen zur Robustheit
- ggf. Untersuchungen zur Übertragbarkeit

Sofern Sie im Dossier Ergebnisse zu Surrogatendpunkten eingereicht haben, benennen Sie die Gründe für die Verwendung von Surrogatendpunkten. Beschreiben Sie, ob und warum die verwendeten Surrogatendpunkte im betrachteten Kontext valide Surrogatendpunkte darstellen bzw. Aussagen zu patientenrelevanten Endpunkten zulassen.

Nicht zutreffend.

¹⁷ Burzykowski T (Ed.): The evaluation of surrogate endpoints. New York: Springer; 2005.

¹⁸ Molenberghs G, Burzykowski T, Alonso A, Assam P, Tilahun A, Buyse M: A unified framework for the evaluation of surrogate endpoints in mental-health clinical trials. Stat Methods Med Res 2010; 19(3): 205-236.

¹⁹ Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. Pharm Stat 2006; 5(3): 173-186.

²⁰ Weir CJ, Walley RJ. Statistical evaluation of biomarkers as surrogate endpoints: a literature review. Stat Med 2006; 25(2): 183-203.

4.6 Referenzliste

Listen Sie nachfolgend alle Quellen (z. B. Publikationen, Studienberichte, Studienregister-einträge), die Sie im vorliegenden Dokument angegeben haben (als fortlaufend nummerierte Liste). Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard). Geben Sie bei Fachinformationen immer den Stand des Dokuments an.

1. Acharya S. und Musa R. 2024. *Polymyalgia Rheumatica*. Last Update: February 25, 2024. Verfügbar unter: <https://www.ncbi.nlm.nih.gov/books/NBK537274>, abgerufen am: 20.01.2025.
2. acis Arzneimittel GmbH (acis) 2024a. *Fachinformation Prednisolon acis: Stand Juni 2024*. Verfügbar unter: www.fachinfo.de, abgerufen am: 20.01.2025.
3. acis Arzneimittel GmbH (acis) 2024b. *Fachinformation Prednison acis: Stand Juni 2024*. Verfügbar unter: www.fachinfo.de, abgerufen am: 20.01.2025.
4. Binard A., Lefebvre B., Bandt M. de et al. 2009. *Validity of the polymyalgia rheumatica activity score in primary care practice*. *Annals of the rheumatic diseases* 68 (4), S. 541–545.
5. BPAC NZ Limited (BPAC) 2023. *Polymyalgia rheumatica (PMR) – look before you leap*. Verfügbar unter: <https://bpac.org.nz/2023/pmr.aspx#:~:text=The%20typical%20presentation%20of%20polymyalgia,50%20years%20of%20European%20ethnicity>, abgerufen am: 16.10.2024.
6. Bridgewater S., Shepherd M. A., Dawson J. et al. 2023. *Measuring the impact of steroid therapy on health-related quality of life in patients with rheumatic diseases: international development of a glucocorticoid treatment-specific patient-reported outcome measure*. *Rheumatology (Oxford, England)* 62 (11), S. 3565–3575.
7. Bruce B. und Fries J. F. 2003. *The Stanford Health Assessment Questionnaire: dimensions and practical applications*. *Health and quality of life outcomes* 1, S. 20.
8. Bullinger M., Kirchberger I. und Ware J. 1995. *Der deutsche SF-36 Health Survey Übersetzung und psychometrische Testung eines krankheitsübergreifenden Instruments zur Erfassung der gesundheitsbezogenen Lebensqualität*. *Zeitschrift für Gesundheitswissenschaften = Journal of public health* 3 (1), S. 21–36.
9. Bundesministerium der Justiz (BMJ) 2023. *Verordnung über die Nutzenbewertung von Arzneimitteln nach § 35a Absatz 1 SGB V für Erstattungsvereinbarungen nach § 130b SGB V (Arzneimittel-Nutzenbewertungsverordnung - AM-NutzenV): Zuletzt geändert durch Art. 6 G v. 19.7.2023 I Nr. 197*. Verfügbar unter: <https://www.gesetze-im-internet.de/am-nutzenv/>, abgerufen am: 29.11.2023.
10. Butt Z., Lai J.-S., Rao D. et al. 2013. *Measurement of fatigue in cancer, stroke, and HIV using the Functional Assessment of Chronic Illness Therapy - Fatigue (FACIT-F) scale*. *Journal of psychosomatic research* 74 (1), S. 64–68.
11. Buttgerit F., Brabant T., Dinges H. et al. 2018. *S3-Leitlinie zur Behandlung der Polymyalgia rheumatica: Evidenzbasierte Leitlinie der Deutschen Gesellschaft für*

- Rheumatologie (DGRh), der Österreichischen Gesellschaft für Rheumatologie und Rehabilitation (ÖGR) und der Schweizerischen Gesellschaft für Rheumatologie (SGR) und der beteiligten medizinisch-wissenschaftlichen Fachgesellschaften und weiterer Organisationen. Zeitschrift für Rheumatologie 77 (5), S. 429–441.*
12. Caporali R., Cimmino M. A., Ferraccioli G. et al. 2004. *Prednisone plus methotrexate for polymyalgia rheumatica: a randomized, double-blind, placebo-controlled trial.* Annals of internal medicine 141 (7), S. 493–500.
 13. Catanoso M. G., Macchioni P., Boiardi L. et al. 2007. *Treatment of refractory polymyalgia rheumatica with etanercept: an open pilot study.* Arthritis and rheumatism 57 (8), S. 1514–1519.
 14. Cella D., Yount S., Sorensen M. et al. 2005. *Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis.* The Journal of rheumatology 32 (5), S. 811–819.
 15. ClinicalTrialsGOV 2022. *A Randomized, Double-blind, Placebo-controlled Study to Evaluate the Efficacy and Safety of Sarilumab in Patients With Polymyalgia Rheumatica.* NCT03600818 / EFC15160 / 2017-002989-42 / U1111-1201-0777. Verfügbar unter: <https://clinicaltrials.gov/study/NCT03600818>, abgerufen am: 22.11.2024.
 16. Dejaco C., Duftner C., Cimmino M. A. et al. 2011. *Definition of remission and relapse in polymyalgia rheumatica: data from a literature search compared with a Delphi-based expert consensus.* Annals of the rheumatic diseases 70 (3), S. 447–453.
 17. Dejaco C., Singh Y. P., Perel P. et al. 2015. *2015 Recommendations for the management of polymyalgia rheumatica: a European League Against Rheumatism/American College of Rheumatology collaborative initiative.* Annals of the rheumatic diseases 74 (10), S. 1799–1807.
 18. Dermapharm 2022. *Fachinformation Volon® 4/8/16 mg: Stand März 2022.* Verfügbar unter: www.fachinfo.de, abgerufen am: 20.01.2025.
 19. Ellert U. und Kurth B. M. 2013. *Gesundheitsbezogene Lebensqualität bei Erwachsenen in Deutschland.* Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz 56 (5), S. 643–649.
 20. EU-CTR 0000. *A randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of sarilumab in patients with polymyalgia rheumatica.* 2017-002989-42 / EFC15160. Verfügbar unter: https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-002989-42, abgerufen am: 22.11.2024.
 21. European Medicines Agency (EMA) 2024. *Assessment Report Kevzara / sarilumab (EPAR: Procedure No. EMEA/H/C/004254/II/0044).* Verfügbar unter: https://www.ema.europa.eu/en/documents/variation-report/kevzara-h-c-004254-ii-0044-epar-assessment-report-variation_en.pdf, abgerufen am: 11.12.2024.
 22. Fries J. F., Spitz P., Kraines R. G. et al. 1980. *Measurement of patient outcome in arthritis.* Arthritis and rheumatism 23 (2), S. 137–145.

23. Gabriel S. E., Sunku J., Salvarani C. et al. 1997. *Adverse outcomes of antiinflammatory therapy among patients with polymyalgia rheumatica*. Arthritis and rheumatism 40 (10), S. 1873–1878.
24. GALENpharma GmbH (Galen) 2024. *Fachinformation Metypred® 125 / 250 / forte 1000 mg: Stand September 2024*. Verfügbar unter: <https://galenpharma.de/produkt/metypred-forte-galen/>, abgerufen am: 20.01.2024.
25. Gemeinsamer Bundesausschuss (G-BA) 2018. *Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM-RL): Anlage XII - Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Sarilumab*. Verfügbar unter: https://www.g-ba.de/downloads/40-268-4803/2018-02-15_AM-RL-XII_Sarilumab_D-299_TrG.pdf, abgerufen am: 29.11.23.
26. Gemeinsamer Bundesausschuss (G-BA) 2021. *Tragende Gründe zum Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Arzneimittel-Richtlinie (AM RL): Anlage XII – Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V Lumasiran (Hyperoxalurie)*. Verfügbar unter: https://www.g-ba.de/downloads/40-268-7656/2021-07-01_AM-RL-XII_Lumasiran_D-622_TrG.pdf, abgerufen am: 05.10.2023.
27. Gemeinsamer Bundesausschuss (G-BA) 2024. *Niederschrift (finale Fassung) zum Beratungsgespräch gemäß § 8 Abs. 1 AM-NutzenV - Beratungsanforderung 2023-B-266: Sarilumab zur Behandlung der rheumatischen Polymyalgie*.
28. Gemeinsamer Bundesausschuss (G-BA) 2025. *Verfahrensordnung des Gemeinsamen Bundesausschusses*. In Kraft getreten am 3. Januar 2025. Verfügbar unter: https://www.g-ba.de/downloads/62-492-3689/VerfO_2024-09-19_iK_2025-01-03.pdf, abgerufen am: 08.01.2025.
29. Herdman M., Gudex C., Lloyd A. et al. 2011. *Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L)*. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 20 (10), S. 1727–1736.
30. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2023. *Allgemeine Methoden Version 7.0 vom 19.09.2023*. Verfügbar unter: https://www.iqwig.de/methoden/allgemeine-methoden_version-7-0.pdf, abgerufen am: 05.10.2023.
31. Janssen M. F., Pickard A. S., Golicki D. et al. 2013. *Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study*. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation 22 (7), S. 1717–1727.
32. Kalke S., Mukerjee D. und Dasgupta B. 2000. *A study of the health assessment questionnaire to evaluate functional status in polymyalgia rheumatica*. Rheumatology (Oxford, England) 39 (8), S. 883–885.

33. Leeb B. F. und Bird H. A. 2004. *A disease activity score for polymyalgia rheumatica*. Annals of the rheumatic diseases 63 (10), S. 1279–1283.
34. Leeb B. F., Rintelen B., Sautner J. et al. 2007. *The polymyalgia rheumatica activity score in daily use: proposal for a definition of remission*. Arthritis and rheumatism 57 (5), S. 810–815.
35. Leung J. L., Twohig H., Muller S. et al. 2023. *Test-retest reliability of pain VAS/NRS, stiffness VAS/NRS, HAQ-DI and mHAQ in polymyalgia rheumatica: An OMERACT study*. Seminars in arthritis and rheumatism 62, S. 152239.
36. Liew D. F., Owen C. E. und Buchanan R. R. 2018. *Prescribing for polymyalgia rheumatica*. Australian prescriber 41 (1), S. 14–19.
37. Lundberg I. E., Sharma A., Turesson C. et al. 2022. *An update on polymyalgia rheumatica*. Journal of internal medicine 292 (5), S. 717–732.
38. Montan I., Löwe B., Cella D. et al. 2018. *General Population Norms for the Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale*. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research 21 (11), S. 1313–1321.
39. Morfeld M. und Bullinger M. 2008. *Der SF-36 Health Survey zur Erhebung und Dokumentation gesundheitsbezogener Lebensqualität*. Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin 18 (05), S. 250–255.
40. Salvarani C., Macchioni P., Manzini C. et al. 2007. *Infliximab plus prednisone or placebo plus prednisone for the initial treatment of polymyalgia rheumatica: a randomized trial*. Annals of internal medicine 146 (9), S. 631–639.
41. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024a. *Fachinformation Kevzara® 150 mg/200 mg Injektionslösung in einer Fertigspritze / Kevzara® 150 mg/200 mg Injektionslösung im Fertigen: Stand November 2024*. Verfügbar unter: www.fachinfo.de, abgerufen am: 14.01.2025.
42. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024b. *Zusatzanalysen zur Studie SAPHYR - Efficacy (Continuous Analyses - Subgroups)*.
43. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024c. *Zusatzanalysen zur Studie SAPHYR - Efficacy (Continuous Analyses)*.
44. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024d. *Zusatzanalysen zur Studie SAPHYR - Efficacy (Responder Analyses - Subgroups)*.
45. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024e. *Zusatzanalysen zur Studie SAPHYR - Efficacy (Responder Analyses)*.
46. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024f. *Zusatzanalysen zur Studie SAPHYR - Efficacy (Time to Event Analyses - Subgroups)*.
47. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024g. *Zusatzanalysen zur Studie SAPHYR - Efficacy (Time to Event Analyses)*.

48. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024h. *Zusatzanalysen zur Studie SAPHYR - Safety.*
49. Sanofi-Aventis Deutschland GmbH (Sanofi) 2024i. *Zusatzanalysen zur Studie SAPHYR - Safety (Subgroup Analyses).*
50. Sanofi-Aventis Recherche & Développement, Frankreich (Sanofi) 2021a. *Clinical Study Report: A randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of sarilumab in patients with polymyalgia rheumatica.*
51. Sanofi-Aventis Recherche & Développement, Frankreich (Sanofi) 2021b. *Clinical Trial Protocol 02: A randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of sarilumab in patients with polymyalgia rheumatica.*
52. Sanofi-Aventis Recherche & Développement, Frankreich (Sanofi) 2021c. *Efficacy Response Data - Appendix 16.2.6, Clinical Study Report: A randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of sarilumab in patients with polymyalgia rheumatica.*
53. Sanofi-Aventis Recherche & Développement (Sanofi) 2021d. *Statistical Analysis Plan: A randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of sarilumab in patients with polymyalgia rheumatica.* Data on File.
54. Spiera R. F., Unizony S., Warrington K. J. et al. 2023. *Sarilumab for Relapse of Polymyalgia Rheumatica during Glucocorticoid Taper.* CN-02601940. The New England journal of medicine 389 (14), S. 1263–1272.
55. Steindl M., Nothnagl T., Andel I. et al. 2009. *Diagnose und Monitoring bei Polymyalgia Rheumatica.* Aktuelle Rheumatologie 34 (03), S. 171–174.
56. Sullivan P. W., Ghushchyan V. H., Globe G. et al. 2017. *Health-related quality of life associated with systemic corticosteroids.* Quality of Life Research 26 (4), S. 1037–1058.
57. Ware J. E., JR 2000. *SF-36 health survey update.* Spine 25 (24), S. 3130–3139.
58. Ware J. E., JR und Sherbourne C. D. 1992. *The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection.* Medical care 30 (6), S. 473–483.
59. Weyand C. M., Fulbright J. W., Evans J. M. et al. 1999. *Corticosteroid requirements in polymyalgia rheumatica.* Archives of internal medicine 159 (6), S. 577–584.
60. WHO-ICTRP 2021. *Evaluation of the Efficacy and Safety of Sarilumab in Patients With Polymyalgia Rheumatica.* NCT03600818 / 2017-002989-42/U1111-1201-0777/EFC15160. Verfügbar unter: <https://trialssearch.who.int/Trial2.aspx?TrialID=NCT03600818>, abgerufen am: 22.11.2024.
61. Wong S. S.-L., Wilczynski N. L. und Haynes R. B. 2006. *Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE.* Journal of the Medical Library Association 94 (4), S. 451–455.

Anhang 4-A: Suchstrategien – bibliografische Literaturrecherche

Geben Sie nachfolgend die Suchstrategien für die bibliografische(n) Literaturrecherche(n) an, und zwar getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.). Für jede durchsuchte Datenbank ist die verwendete Strategie separat darzustellen. Geben Sie dabei zunächst jeweils den Namen der durchsuchten Datenbank (z. B. EMBASE), die verwendete Suchoberfläche (z. B. DIMDI, Ovid etc.), das Datum der Suche, das Zeitsegment (z. B.: „1980 to 2010 week 50“) und die gegebenenfalls verwendeten Suchfilter (mit Angabe einer Quelle) an. Listen Sie danach die Suchstrategie einschließlich der resultierenden Trefferzahlen auf. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel (eine umfassende Suche soll Freitextbegriffe und Schlagwörter enthalten):

Datenbankname	EMBASE	
Suchoberfläche	Ovid	
Datum der Suche	07.11.2016	
Zeitsegment	1974 to 2016 November 04	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 [Quelle ²¹] – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
1	Diabetes Mellitus/	552986
2	Non Insulin Dependent Diabetes Mellitus/	195234
3	(diabet* or niddm or t2dm).ab,ti.	714228
4	or/1-3	847068
5	linagliptin*.mp.	1562
6	(random* or double-blind*).tw.	1193849
7	placebo*.mp.	388057
8	or/6-7	1382838
9	and/4,5,8	633

²¹ Das Zitat zu dem hier beispielhaft angegebenen Suchfilter lautet wie folgt: Wong SSL, Wilczynski NL, Haynes RB. Comparison of top-performing search strategies for detecting clinically sound treatment studies and systematic reviews in MEDLINE and EMBASE. J Med Libr Assoc 2006; 94(4): 451-455. Hinweis: Für die Suche in der Cochrane-Datenbank „Cochrane Central Register of Controlled Trials (Clinical Trials)“ sollte kein Studienfilter verwendet werden.

Anhang 4-A1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Datenbankname	Ovid MEDLINE® and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations and Daily	
Suchoberfläche	Ovid	
Datum der Suche	22.11.2024	
Zeitsegment	1946 to November 21, 2023	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp Polymyalgia Rheumatica/	2.916
2	((polymyalg* and rheuma*) or PMR).mp.	6.826
3	1 or 2	6.826
Intervention		
4	(Sarilumab* or SAR 153191 or SAR?153191 or REGN?88 or Kevzara*).mp.	404
Studientyp		
5	randomized controlled trial.pt. or randomized.mp. or placebo.mp.	1.176.494
Zusammenführung		
6	3 and 4 and 5	2
7	remove duplicates from 6	2

Datenbankname	Embase 1974 to 2024 November 21	
Suchoberfläche	Ovid	
Datum der Suche	22.11.2024	
Zeitsegment	1974 to 2024 November 21	
Suchfilter	Filter für randomisierte kontrollierte Studien nach Wong 2006 – Strategy minimizing difference between sensitivity and specificity	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp rheumatic polymyalgia/	7.404
2	((polymyalg* and rheuma*) or PMR).mp.	13.023
3	1 or 2	13.023
Intervention		
4	exp Sarilumab/	2.392
5	(Sarilumab* or SAR 153191 or SAR?153191 or REGN88 or REGN-88 or Kevzara*).mp.	2.456
6	4 or 5	2.456
Studientyp		
7	random*.tw. or placebo*.mp. or double-blind*.tw.	2.431.266
Zusammenführung		
8	3 and 6 and 7	14
9	remove duplicates from 8	14

Datenbankname	EBM Reviews - Cochrane Central Register of Controlled Trials	
Suchoberfläche	Ovid	
Datum der Suche	22.11.2024	
Zeitsegment	October 2024	
Suchfilter	-	
#	Suchbegriffe	Ergebnis
Indikation		
1	exp Polymyalgia Rheumatica/	125
2	((polymyalg* and rheuma*) or PMR).mp.	897
3	1 or 2	897
Intervention		
4	(Sarilumab* or SAR 153191 or SAR?153191 or REGN?88 or Kevzara*).mp.	358
Zusammenführung		
5	3 and 4	10
6	remove duplicates from 5	10

Anhang 4-A2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-A3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-A4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-B: Suchstrategien – Suche in Studienregistern/ Studienergebnisdatenbanken

Geben Sie nachfolgend die Suchstrategien für die Suche(n) in Studienregistern/ Studienergebnisdatenbanken an. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Für jede/s durchsuchte Studienregister/ Studienergebnisdatenbank ist eine separate Strategie darzustellen. Geben Sie dabei jeweils den Namen des durchsuchten Studienregisters/ Studienergebnisdatenbank (z. B. [clinicaltrials.gov](http://www.clinicaltrials.gov)), die Internetadresse, unter der das/die Studienregister/ Studienergebnisdatenbank erreichbar ist (z. B. <http://www.clinicaltrials.gov>), das Datum der Suche, die verwendete Suchstrategie und die resultierenden Treffer an. Orientieren Sie sich bei Ihren Angaben an dem nachfolgenden Beispiel:

Studienregister/ Studienergebnisdatenbank	International Clinical Trials Registry Platform Search Portal
Internetadresse	http://apps.who.int/trialsearch/
Datum der Suche	07.11.2016
Eingabeoberfläche	Standard Search
Suchstrategie	linagliptin OR BI 1356
Treffer	169

Anhang 4-B1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Studienregister	clinicaltrials.gov
Internetadresse	https://clinicaltrials.gov/
Datum der Suche	22.11.2024
Eingabeoberfläche	Standard search – other terms
Suchstrategie	(Polymyalgia Rheumatica OR PMR) AND (sarilumab OR SAR153191 OR SAR-153191 OR “SAR 153191” OR REGN88 OR REGN-88 OR “REGN 88” OR Kevzara)
Treffer	2

Studienregister	EU Clinical Trials Register
Internetadresse	https://www.clinicaltrialsregister.eu/
Datum der Suche	22.11.2024
Eingabeoberfläche	Standard search
Suchstrategie	("Polymyalgia Rheumatica" OR PMR) AND (sarilumab OR SAR153191 OR SAR-153191 OR “SAR 153191” OR REGN88 OR REGN-88 OR “REGN 88” OR Kevzara)
Treffer	2

Studienregister	ICTRP (WHO)
Internetadresse	http://apps.who.int/trialsearch/Default.aspx
Datum der Suche	22.11.2024
Eingabeoberfläche	Standard search
Suchstrategie	(“Polymyalgia Rheumatica” OR PMR) AND (sarilumab OR SAR153191 OR SAR-153191 OR “SAR 153191” OR REGN88 OR REGN-88 OR “REGN 88” OR Kevzara)
Treffer	1

Anhang 4-B2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-B3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-B4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-C: Liste der im Volltext gesichteten und ausgeschlossenen Dokumente mit Ausschlussgrund (bibliografische Literaturrecherche)

Listen Sie nachfolgend die im Volltext gesichteten und ausgeschlossenen Dokumente aus der /den bibliografischen Literaturrecherche(n) auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-C1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Es wurden keine der im Volltext gesichteten Dokumente ausgeschlossen.

Anhang 4-C2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-C3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-C4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-D: Liste der ausgeschlossenen Studien mit Ausschlussgrund (Suche in Studienregistern/ Studienergebnisdatenbanken)

Listen Sie nachfolgend die durch die Studienregistersuche(n)/ Studienergebnisdatenbanksuche(n) identifizierten, aber ausgeschlossenen Registereinträge auf. Machen Sie die Angaben getrennt für die einzelnen Recherchen (Suche nach RCT mit dem zu bewertenden Arzneimittel, Suche nach RCT für indirekte Vergleiche etc.) wie unten angegeben. Verwenden Sie hierzu einen allgemein gebräuchlichen Zitierstil (z. B. Vancouver oder Harvard) und nummerieren Sie die Zitate fortlaufend. Geben Sie jeweils einen Ausschlussgrund an und beziehen Sie sich dabei auf die im Abschnitt 4.2.2 genannten Ein- und Ausschlusskriterien.

Anhang 4-D1: Suche nach RCT mit dem zu bewertenden Arzneimittel

Nr.	Studienbezeichnung	Studienbeschreibung	Ausschlussgrund
1	NCT03600805 / EFC15068 / 2017-002988-18 / U1111-1200-2184	A Randomized, Double-blind, Placebo-controlled Study to Evaluate the Efficacy and Safety of Sarilumab in Patients With Giant Cell Arteritis. [Zugriff: 22.11.2024] https://clinicaltrials.gov/study/NCT03600805	A1 - Population
2	2017-002988-18 / EFC15068	A randomized, double-blind, placebo-controlled study to evaluate the efficacy and safety of sarilumab in patients with giant cell arteritis. [Zugriff: 22.11.2024] https://www.clinicaltrialsregister.eu/ctr-search/search?query=eudract_number:2017-002988-18	A1 - Population

Anhang 4-D2: Suche nach RCT für indirekte Vergleiche

Nicht zutreffend.

Anhang 4-D3: Suche nach nicht randomisierten vergleichenden Studien

Nicht zutreffend.

Anhang 4-D4: Suche nach weiteren Untersuchungen

Nicht zutreffend.

Anhang 4-E: Methodik der eingeschlossenen Studien – RCT

Beschreiben Sie nachfolgend die Methodik jeder eingeschlossenen, in Abschnitt 4.3.1.1.5 genannten Studie. Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-94 inklusive eines Flow-Charts für den Patientenfluss.

Sollten Sie im Dossier indirekte Vergleiche präsentieren, beschreiben Sie ebenfalls die Methodik jeder zusätzlich in den indirekten Vergleich eingeschlossenen Studie (Abschnitt 4.3.2.1). Erstellen Sie hierfür je Studie eine separate Version der nachfolgend dargestellten Tabelle 4-94 inklusive eines Flow-Charts für den Patientenfluss.

Tabelle 4-94 (Anhang): Studiendesign und -methodik für Studie SAPHYR

Item ^a	Charakteristikum	Studieninformation
Studienziel		
2b	Genauere Ziele, Fragestellung und Hypothesen	<p><u>Fragestellung:</u> Bewertung der Wirksamkeit und Sicherheit von Sarilumab bei Patienten mit aktiver Polymyalgia rheumatica (PMR).</p> <p><u>Primäres Ziel:</u></p> <ul style="list-style-type: none"> • Bewertung der Wirksamkeit von Sarilumab bei Patienten mit PMR anhand des Anteils an Patienten mit anhaltender Remission zu Woche 52 mit 14-wöchigem Ausschleichen der CS, im Vergleich zu Placebo mit 52-wöchigem CS-Ausschleichen. <p><u>Sekundäre Ziele:</u></p> <ul style="list-style-type: none"> • Nachweis der Wirksamkeit von Sarilumab mit 14-wöchigem CS-Ausschleichen bei Patienten mit PMR in Bezug auf: <ul style="list-style-type: none"> ○ Klinisches Ansprechen (z. B. anhaltende Remission, Remissionsraten, Zeit bis zum ersten Krankheitsschub) im Zeitverlauf ○ Kumulativer CS-Gebrauch (einschließlich Prednison) • Bewertung der Sicherheit (einschließlich Immunogenität) und Verträglichkeit von Sarilumab bei Patienten mit PMR • Messung der Serumkonzentrationen von Sarilumab bei Patienten mit PMR • Bewertung der Wirkung von Sarilumab auf die Verringerung der Glukokortikoid-Toxizität, gemessen anhand des zusammengesetzten GTI. <p><u>Weitere Ziele:</u></p> <ul style="list-style-type: none"> • Bewertung der Wirkung von Sarilumab auf die Prüfarzt-Beurteilung der Krankheitsaktivität mittels VAS (Physician Global Assessment of Disease Activity – Visual Analog Scale) • Bewertung der Wirkung von Sarilumab auf PRO, einschließlich <ul style="list-style-type: none"> ○ Fatigue (anhand der Ermüdungsskala des funktionellen Beurteilungsinstruments zur Therapie chronischer Erkrankungen (FACIT-F)) ○ Gesundheitsstatus (anhand der VAS des European Quality of Life 5 Dimensions-Fragebogens, EQ-5D VAS, und des 36-Item Short-Form Health Survey-Kurzfragebogens Version 2, SF-36v2) ○ Patientenbeurteilung des Schmerzempfindens (anhand des HAQ-DI über eine VAS) ○ Patientenbeurteilung der Krankheitsaktivität (anhand des HAQ-DI über eine VAS) • Bewertung des Einflusses der Level von ESR und CRP auf den Remissionsstatus • Charakterisierung der Krankheitsaktivität bei Patienten mit PMR während des CS-Ausschleichens oder Sarilumab-

Item ^a	Charakteristikum	Studieninformation
		<p>Behandlung bei einer Teilpopulation der Patienten mit Hilfe der zirkulierenden Immuzelltypen</p> <ul style="list-style-type: none"> Charakterisierung der Krankheitsaktivität bei Patienten mit PMR während des CS-Ausschleichens oder Sarilumab-Behandlung anhand von zirkulierenden Proteinen, Genetik und Genexpression bei Patienten, die ihre Einwilligung zu diesem optionalen Studienteil gegeben haben. <p><u>Hypothese:</u> Sarilumab (+ 14-wöchiges Prednison-Ausschleichen) ist Placebo (+ 52-wöchiges Prednison-Ausschleichen) hinsichtlich des Anteils an Patienten mit anhaltender Remission zu Woche 52 überlegen.</p>
Methoden		
3	Studiendesign	
3a	Beschreibung des Studiendesigns (z. B. parallel, faktoriell) inklusive Zuteilungsverhältnis	<p>Multizentrische, randomisierte, doppelblinde, Placebo-kontrollierte, 52-wöchige Phase III-Studie mit zwei parallelen Behandlungsgruppen</p> <p>Zuteilungsverhältnis 1:1 (Sarilumab:Placebo)</p>
3b	Relevante Änderungen der Methodik nach Studienbeginn (z. B. Ein-/Ausschlusskriterien), mit Begründung	<p>Es wurden zwei Amendments zum Original-Studienprotokoll vom 08. Februar 2018 erstellt. Das erste wurde jedoch bereits vor Einschluss des ersten Patienten verabschiedet. Entsprechend gab es ein Amendment (02) nach Studienbeginn mit relevanten Änderungen der Methodik:</p> <p><u>Amendment 02 (19. April 2021):</u></p> <p>Aufgrund der Rekrutierungsproblematik während der COVID-19-Pandemie:</p> <ul style="list-style-type: none"> Reduzierung der geplanten Patientenzahl von 280 auf 118 Änderung der statistischen Signifikanzschwelle für den primären und die sekundären Wirksamkeitsendpunkte von 0,01 auf 0,05 Änderung des KI von 99 % auf 95 %
4	Probanden / Patienten	
4a	Ein-/Ausschlusskriterien der Probanden / Patienten	<p><u>Einschlusskriterien:</u></p> <ul style="list-style-type: none"> Schriftliche Einwilligungserklärung PMR-Diagnose anhand der Klassifikationskriterien der EULAR: <ul style="list-style-type: none"> Alter \geq 50 Jahre zur Zeit der Diagnose Beidseitige Schulterschmerzen Erhöhte Akut-Phase-Reaktanten (CRP > 10 mg/l und / oder ESR > 30 mm/h) <p>UND eines der folgenden:</p> <ul style="list-style-type: none"> Punktwert \geq 4 zur Zeit der Diagnose auf Basis von (ohne Ultraschall): <ul style="list-style-type: none"> - Dauer der Morgensteifigkeit > 45 Minuten: 2 Punkte

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> - Hüftschmerzen oder eingeschränkte Beweglichkeit: 1 Punkt - Seronegativ für Rheumafaktor (RF) und anti-zyklisches citrulliniertes Peptid (anti-CCP): 2 Punkte - Keine andere Gelenkbeteiligung: 1 Punkt <p>ODER</p> <ul style="list-style-type: none"> ○ Punktwert ≥ 5 zur Zeit der Diagnose auf Basis von (mit Ultraschall): <ul style="list-style-type: none"> - Dauer der Morgensteifigkeit > 45 Minuten: 2 Punkte - Hüftschmerzen oder eingeschränkte Beweglichkeit: 1 Punkt - Seronegativ für RF und anti-CCP: 2 Punkte - Keine andere Gelenkbeteiligung: 1 Punkt - Mindestens eine Schulter mit Schleimbeutelentzündung und / oder Sehnenscheidenentzündung des Bizeps und / oder Synovitis des Schulterhauptgelenks (posterior oder axillär) und mindestens eine Hüfte mit Synovitis und / oder Schleimbeutelentzündung am Trochanter: 1 Punkt (durch Ultraschall bestätigt) - Beide Schultern mit Schleimbeutelentzündung, Sehnenscheidenentzündung des Bizeps oder Synovitis des Schulterhauptgelenks: 1 Punkt (durch Ultraschall bestätigt) • Patient unter Prednison-Behandlung von mindestens 7,5 mg/Tag (oder Äquivalent) und nicht mehr als 20 mg/Tag zu Screening und während der Screening-Phase • Patient willens und in der Lage, Prednison in einer Dosis von 15 mg/Tag zur Randomisierung zu erhalten • Patient in der Vorgeschichte über mindestens acht Wochen mit Prednison in einer Dosis von ≥ 10 mg/Tag oder Äquivalent behandelt • Patient mit mindestens einem eindeutigen PMR-Schub innerhalb der letzten 12 Wochen vor Screening während des Versuchs, Prednison bei einer Dosis von $\geq 7,5$ mg/Tag (oder Äquivalent) auszuschleichen <ul style="list-style-type: none"> ○ Eindeutige PMR-Symptome wurden definiert als Schmerzen in Schulter und / oder Hüftgürtel, verbunden mit Entzündungssteifheit • Patient mit ESR ≥ 30 mm/h oder CRP ≥ 10 mg/l, verbunden mit PMR-Krankheitsaktivität innerhalb von 12 Wochen vor Screening <p><u>Ausschlusskriterien:</u></p> <ul style="list-style-type: none"> • Diagnose einer Riesenzellarteriitis (z. B. persistierende oder wiederkehrende lokalisierte Kopfschmerzen, Empfindlichkeit der Temporalarterie oder der Kopfhaut, Bewegungsschmerz in Kiefer und / oder Extremitäten,

Item ^a	Charakteristikum	Studieninformation
		<p>verschwommenes Sehen oder Sehverlust, Schlafanfall-Symptome)</p> <ul style="list-style-type: none"> • Gleichzeitige rheumatoide Arthritis oder andere entzündliche Arthritis oder andere Bindegewebserkrankungen, z. B. systemischer Lupus erythematodes, systemische Sklerose, Vaskulitis, Myositis, gemischte Bindegewebserkrankung, ankylosierende Spondylitis • Gleichzeitige Diagnose einer Rhabdomyolyse oder neuropathischen muskulären Erkrankung • Gleichzeitige Diagnose einer aktiven Fibromyalgie • Unzureichend behandelte Hypothyreose • Organtransplantiertes Patient • Jede vorherige (innerhalb der nachfolgend definierten Zeitspannen) oder gleichzeitige Behandlung mit Immunsuppressiva, z. B.: <ul style="list-style-type: none"> ○ Janus-Kinase-Inhibitor (z. B. Tofacitinib) innerhalb von vier Wochen vor Baseline ○ Wirkstoffe zur Zelldepletion (z. B. Anti-CD20) ohne erkennbare Erholung der B-Zellen auf Baseline-Niveau ○ Anakinra innerhalb einer Woche vor Baseline ○ Abatacept innerhalb von acht Wochen vor Baseline ○ Tumor-Nekrose-Faktor-Inhibitoren innerhalb von zwei bis acht Wochen (Etanercept innerhalb von zwei Wochen, Infliximab, Certolizumab, Golimumab oder Adalimumab innerhalb von acht Wochen) oder nach Ablauf von mindestens fünf Halbwertszeiten, je nachdem, was länger war ○ Alkylierungsmittel, einschließlich Cyclophosphamid innerhalb von sechs Monaten vor Baseline ○ Cyclosporin, Azathioprin oder Mycophenolat-Mofetil oder Leflunomid innerhalb von vier Wochen vor Baseline • Therapieversagen, einschließlich unzureichendem Ansprechen oder Intoleranz oder Kontraindikation auf / gegenüber biologischer IL-6-Antagonisten (vorherige Behandlung mit IL-6-Antagonisten, die nicht aufgrund von Therapieversagen mindestens drei Monate vor Baseline beendet wurde, war nicht ausgeschlossen) • Instabile MTX-Dosis und / oder MTX-Dosis > 15 mg/Woche innerhalb von drei Monaten vor Baseline • Gleichzeitige Behandlung mit CS in anderer Indikation als PMR • Teilnahme an einer klinischen Medikamentenstudie innerhalb von fünf Halbwertszeiten oder 60 Tagen vor der Screening-Visite, je nachdem, was länger war • Vorgeschichte eines Alkohol- oder Drogenmissbrauchs innerhalb von fünf Jahren vor der Screening-Visite

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Patient, der die Zustimmung zur Studienteilnahme während der Screening-Phase (nach Unterzeichnung der Einwilligungserklärung) widerrief • Nicht fähig oder nicht willens, die PRO-Fragebögen auszufüllen • Jede der folgenden Zustände / Situationen bei einem Patienten: <ul style="list-style-type: none"> ○ Kurze Lebenserwartung ○ Zustände / Begleiterkrankungen, die die Bewertung der Wirksamkeitsendpunkte verhinderten (z. B. Patienten mit chronischen Schmerzen) ○ Patient war Prüfarzt, Co-Prüfarzt, Forschungsassistent, Pharmazeut, Studienkoordinator, anderes Mitglied des Studienteams oder dessen Angehöriger, unmittelbar involviert in die Studiendurchführung, oder, falls zutreffend, Mitarbeiter des Prüfzentrums / Prüfarztes oder des Sponsors ○ Unkooperatives oder anderes Verhalten, das die Einhaltung der Studienprozeduren gefährden konnte; Personen, die sich auf Anordnung der Justiz- oder Verwaltungsbehörden in einer Institution befanden • Schwangere oder stillende Frauen • Frauen im gebärfähigen Alter, die nicht durch eine hochwirksame Kontrazeptionsmethode geschützt waren, und / oder die einem Schwangerschaftstest nicht zustimmen wollten oder konnten • Tuberkulose (TB)-Ausschlusskriterien: <ul style="list-style-type: none"> ○ Aktive TB oder Vorgeschichte einer unvollständig behandelten TB, unabhängig vom Quantiferon-Ergebnis zu Screening ○ Quantiferon-positive Patienten (ohne aktive Erkrankung) waren von der Studienteilnahme ausgeschlossen mit folgenden Ausnahmen: <ul style="list-style-type: none"> - Patienten mit Vorgeschichte einer dokumentierten abgeschlossenen Chemoprophylaxe einer latenten oder Behandlung einer aktiven TB-Infektion, nach Spezialisten-Konsultation zum Ausschluss einer aktiven TB-Infektion oder zur weiteren Behandlung - Patienten ohne Vorgeschichte einer Chemoprophylaxe einer latenten oder Behandlung einer aktiven TB-Infektion, aber nach Spezialisten-Konsultation zur Initiierung einer geeigneten Chemoprophylaxe entsprechend lokaler Epidemiologie und zutreffender Leitlinien sowie zuverlässige und tolerierte Behandlung seit mindestens einem Monat - Beratung und Zustimmung des Sponsors für jedes der vorstehenden Szenarien erforderlich ○ Klinisch bedeutsame Anomalie, assoziiert mit vorheriger / aktiver TB-Infektion, im Röntgen-Thorax

Item ^a	Charakteristikum	Studieninformation
		<p>mit mindestens posterior-anteriorer Ansicht; zusätzliche laterale Ansicht empfohlen, aber nicht verpflichtend</p> <ul style="list-style-type: none"> ○ Verdacht auf extra-pulmonale TB-Infektion, unabhängig vom Quantiferon-Ergebnis zu Screening ○ Patienten mit hohem Risiko, sich mit TB zu infizieren ○ Patienten nach Impfung mit dem Calmette-Guerin-Bazillus innerhalb von 12 Monaten vor Screening • Patienten mit Vorgeschichte invasiver opportunistischer Infektionen, einschließlich Histoplasmose, Listeriose, Kokzidioidomykose, Candidiasis, Pneumocystis jirovecii, Aspergillose, trotz Heilung, oder John-Cunningham-Virus (progressive multifokale Leukenzephalopathie) • Patienten mit infektiösem Fieber (> 38°C) oder chronische, persistierende oder wiederkehrende Infektion(en) mit Behandlungsbedarf mit Antibiotika, Virostatika oder Antimykotika innerhalb von vier Wochen vor der Screening-Visite, oder andere häufige wiederkehrende Infektionen, vom Prüfarzt als inakzeptabel eingestuft • Patienten mit unkontrollierter Diabetes mellitus, definiert als glykiertes Hämoglobin (HbA1c) $\geq 9\%$ zur Screening-Visite • Patienten mit nicht verheilten oder heilenden Hautgeschwüren • Patienten nach Erhalt abgeschwächter Lebendimpfstoffe innerhalb von drei Monaten vor der Baseline-Visite, z. B. Varicella-Zoster, Schluckimpfungen gegen Polio oder Röteln • Patienten mit positivem Test auf das Hepatitis-B-Oberflächenantigen (HBsAg) oder positivem Test auf Hepatitis-B-Core-Antikörper (HBcAb) mit negativem HBsAg-Test oder Patienten mit positivem Test sowohl auf HBsAg als auch HBcAb bei vorhandener HBV-DNA zu Screening • Patienten mit positivem Test auf Hepatitis-C-Antikörper • Patienten mit positivem Test auf humane Immundefizienz-Virus (HIV)-Antikörper oder mit Verdacht HIV-positiv zu sein • Patienten mit wiederkehrendem Herpes Zoster oder aktivem Herpes Zoster in der Vorgeschichte • Patienten mit Infektion eines Gelenks oder einer Gelenkprothese in der Vorgeschichte • Frühere oder akute bösartige Erkrankungen, einschließlich lymphoproliferativer Erkrankungen, mit Ausnahme von adäquat behandeltem Zervixkarzinom in-situ, nicht metastasiertem Plattenepithel- oder Basalzellkarzinom der Haut, innerhalb von fünf Jahren vor der Baseline-Visite • Frühere oder akute andere bedeutsame Begleiterkrankung(en), die nach Einschätzung des Prüfarztes die Studienteilnahme des Patienten beeinträchtigen würde. Dazu zählten kardiovaskuläre, renale, neurologische, aktive

Item ^a	Charakteristikum	Studieninformation
		<p>infektiöse, endokrinologische, gastrointestinale, hepatobiliäre, metabolische, pulmonale, nicht bösartige lymphoproliferative oder andere lymphatische Erkrankungen</p> <ul style="list-style-type: none"> • Operation innerhalb von vier Wochen vor der Screening-Visite oder während des Studienzeitraums geplante Operation • Patienten mit Vorgeschichte einer systemischen Überempfindlichkeitsreaktion, mit Ausnahme einer lokalen Reaktion an der Einstichstelle, auf ein Biologikum und bekannte Überempfindlichkeit gegen einen Bestandteil von Sarilumab • Patienten mit einer der folgenden Laborauffälligkeiten zur Screening-Visite: <ul style="list-style-type: none"> ○ Hämoglobin < 8,5 g/dl ○ Leukozyten < 3.000/mm³ ○ Neutrophile < 2.000/mm³ ○ Thrombozyten < 150.000 Zellen/mm³ ○ Aspartat-Aminotransferase oder Alanin-Aminotransferase > 1,5 x ULN ○ Gesamtbilirubin > ULN, Ausnahme durch Gentest nachgewiesenes, dokumentiertes Gilbert-Syndrom ○ Schwere unkontrollierte Hypercholesterinämie (> 350 mg/dl) oder Hypertriglyzeridämie (> 500 mg/dl) ○ Berechnete Kreatinin-Clearance < 30 ml/min (anhand Cockcroft-Gault-Formel) • Patienten mit Vorgeschichte einer entzündlichen Darmerkrankung oder schweren Divertikulitis oder Magen-Darm-Perforation
4b	Studienorganisation und Ort der Studiendurchführung	<p>Sponsor der Studie war Sanofi-Aventis, Frankreich.</p> <p>Die Studie wurde an 70 Zentren in 17 Ländern (Australien, Argentinien, Belgien, Deutschland, Estland, Frankreich, Israel, Italien, Japan, Kanada, Niederlande, Russland, Schweiz, Spanien, Ungarn, Vereinigtes Königreich und Vereinigte Staaten von Amerika) durchgeführt.</p>
5	Interventionen Präzise Angaben zu den geplanten Interventionen jeder Gruppe und zur Administration etc.	<p><u>Sarilumab:</u></p> <p>Die Patienten erhielten 200 mg Sarilumab s.c. alle zwei Wochen (q2w) für 52 Wochen in Kombination mit Prednison bzw. einer Kombination aus Prednison und Placebo zu Prednison (während Woche 0 bis Woche 13 einmal täglich, oral) oder Placebo zu Prednison (von Woche 14 bis Woche 51 einmal täglich, oral). Die Prednison-Dosis wurde von 15 mg/Tag in Woche 0 und 1 auf 1 mg/Tag in Woche 13 ausgeschlichen.</p> <p><u>Placebo:</u></p> <p>Die Patienten erhielten Placebo zu Sarilumab s.c. q2w für 52 Wochen in Kombination mit Prednison oder einer Kombination aus Prednison und Placebo zu Prednison für 52 Wochen (einmal täglich, oral). Die Prednison-Dosis wurde von 15 mg/Tag in Woche 0 und 1 auf 1 mg/Tag ab Woche 44 bis Woche 51 ausgeschlichen.</p>

Item ^a	Charakteristikum	Studieninformation
6	Zielkriterien	
6a	Klar definierte primäre und sekundäre Zielkriterien, Erhebungszeitpunkte, ggf. alle zur Optimierung der Ergebnisqualität verwendeten Erhebungsmethoden (z. B. Mehrfachbeobachtungen, Training der Prüfer) und ggf. Angaben zur Validierung von Erhebungsinstrumenten	<p><u>Primärer Wirksamkeitsendpunkt:</u></p> <p>Anteil der Patienten mit anhaltender Remission zu Woche 52; anhaltende Remission definiert als:</p> <ul style="list-style-type: none"> • Erreichen der Remission nicht später als zu Woche 12 UND • Kein Krankheitsschub von Woche 12 bis Woche 52 UND • Anhaltende Reduktion des CRP (auf < 10 mg/l, ohne nachfolgende Anstiege auf ≥ 10 mg/l) von Woche 12 bis Woche 52 UND • Erfolgreiches Einhalten des Prednison-Ausschleichens von Woche 12 bis Woche 52 <p>Krankheits-Remission wurde definiert als Auflösung der Zeichen und Symptome der PMR und CRP-Normalisierung (< 10 mg/l).</p> <p>Die Beurteilung der Zeichen und Symptome der PMR zu jeder Studiervisite sollte unter anderem beinhalten:</p> <ul style="list-style-type: none"> ○ Morgensteifigkeit und / oder Schmerzen in Nacken, Schulter und / oder Hüftgürtel ○ Eingeschränkte Bewegungsfähigkeit in Schultern und / oder Hüftgürtel ○ Konstitutionelle Symptome wie Fatigue, Gewichtsverlust und leichtes Fieber ○ Andere Auffälligkeiten, die nach Ansicht des Prüfarztes auf einen PMR-Schub hinweisen. <p>Krankheitsschub wurde definiert als Wiederauftreten von Zeichen und Symptomen, oder als Anstieg der ESR, im Zusammenhang mit aktiver PMR, die eine Erhöhung der CS-Dosis erforderlich machten. Erhöhung der CS-Dosis wurde definiert als:</p> <ul style="list-style-type: none"> ○ Jede Dosiserhöhung während des protokollgemäßen Ausschleichens ○ Erneute Prednison-Gabe nach Beendigung des protokollgemäßen Ausschleichens <p><u>Sekundäre Endpunkte:</u></p> <p><u>Wirksamkeit:</u></p> <ul style="list-style-type: none"> • Zusammenfassung der Komponenten der anhaltenden Remission zu Woche 52 • Kumulative Prednison-Gesamtdosis (oder Äquivalent) während 52 Wochen • Zeit bis zum ersten PMR-Schub • Veränderung des GTI (und seiner Komponenten) gegenüber Baseline <p><u>Sicherheit:</u></p> <ul style="list-style-type: none"> • UE

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Laborparameter Röntgenthorax • Vitalparameter • körperliche Untersuchung • Elektrokardiogramm <p><u>Andere Wirksamkeitsparameter:</u></p> <ul style="list-style-type: none"> • PRO <ul style="list-style-type: none"> ○ FACIT-Fatigue ○ EQ-5D-3L ○ SF-36v2 ○ HAQ-DI • Prüfarzt-Beurteilung der Krankheitsaktivität mittels MD-VAS <p><u>Pharmakodynamik</u></p> <p><u>Pharmakokinetik</u></p> <p><u>Immunogenität</u></p>
6b	Änderungen der Zielkriterien nach Studienbeginn, mit Begründung	Es gab keine Änderungen der gemäß Studienprotokoll geplanten Analysen der Wirksamkeitsparameter.
7	Fallzahl	
7a	Wie wurden die Fallzahlen bestimmt?	<p>Die Fallzahl wurde aufgrund der Rekrutierungsproblematik während der COVID-19-Pandemie von ursprünglich geplanten 280 auf 118 randomisierte Patienten reduziert.</p> <p>Bei einer Fallzahl von 59 Patienten pro Gruppe liegt die Power bei mindestens 85 % bzw. 95 %, um bei einer angenommenen Ansprechraten zwischen 5 % und 15 % in der Placebo-Gruppe einen Unterschied von 25 % bzw. 30 % der Ansprechraten zwischen Sarilumab (+ 14-wöchiges Prednison-Ausschleichen) und Placebo (+ 52-wöchiges Prednison-Ausschleichen) mittels zweiseitigen χ^2-Tests bei einem Signifikanzniveau von 0,05 aufzuzeigen.</p>
7b	Falls notwendig, Beschreibung von Zwischenanalysen und Kriterien für einen vorzeitigen Studienabbruch	<p>Es waren keine Zwischenanalysen geplant.</p> <p>Der Sponsor konnte die Studie bei unzureichender Patientenaufnahme frühzeitig beenden. Die Rekrutierungsproblematik während der COVID-19-Pandemie bedingte das frühzeitige Studienende.</p>
8	Randomisierung, Erzeugung der Behandlungsfolge	
8a	Methode zur Generierung der zufälligen Zuteilung	Die Randomisierung erfolgte zentral mit Hilfe IRT.
8b	Einzelheiten (z. B. Blockrandomisierung, Stratifizierung)	Die Patienten, die die Einschlusskriterien erfüllten, wurden im Verhältnis 1:1 einer der beiden Behandlungsgruppen zugeteilt.

Item ^a	Charakteristikum	Studieninformation
9	Randomisierung, Geheimhaltung der Behandlungsfolge (allocation concealment) Durchführung der Zuteilung (z. B. nummerierte Behälter; zentrale Randomisierung per Fax / Telefon), Angabe, ob Geheimhaltung bis zur Zuteilung gewährleistet war	Zur Screening-Visite wurde jedem Patienten, der seine Einwilligung zur Studienteilnahme erteilt hatte, via IRT eine zentrumsspezifische Patientennummer in chronologischer Reihenfolge zugeteilt. Zur Baseline-Visite sowie jeder darauffolgenden Visite wurde jedem Patienten, der die Studieneinschlusskriterien erfüllte, via IRT die Nummer des Studienmedikations-Kits zugeteilt. Das Studienzentrum erhielt anschließend eine Bestätigung per Fax bzw. E-Mail.
10	Randomisierung, Durchführung Wer hat die Randomisierungsliste erstellt, wer nahm die Probanden/Patienten in die Studie auf und wer teilte die Probanden/Patienten den Gruppen zu?	Die Randomisierungsliste der Studienmedikationen wurde zentral vom Sponsor der Studie (Sanofi) erstellt. Die Randomisierungsliste der Patienten wurde via IRT erstellt. Die Zuteilung der Patienten zu Sarilumab (+ 14-wöchiges Prednison-Ausschleichen) bzw. Placebo (+ 52-wöchiges Prednison-Ausschleichen) erfolgte über das IRT.
11	Verblindung	
11a	Waren a) die Probanden / Patienten und / oder b) diejenigen, die die Intervention / Behandlung durchführten, und / oder c) diejenigen, die die Zielgrößen beurteilten, verblindet oder nicht verblindet, wie wurde die Verblindung vorgenommen?	Sowohl Prüfarzte als auch Patienten waren bzgl. Verabreichung von Sarilumab oder Placebo sowie des Ausschleichens von Prednison verblindet. Die Prüfarzte hatten ausschließlich in medizinischen Notfallsituationen Zugriff auf die Random-Codes der Studienmedikation.
11b	Falls relevant, Beschreibung der Ähnlichkeit von Interventionen	Sarilumab und das zugehörige Placebo waren nicht unterscheidbar. Dies galt auch für Prednison und das zugehörige Placebo. Die Kits für Sarilumab und das zugehörige Placebo wurden in gleich aussehenden vorgefüllten Glasspritzen zur Verfügung gestellt. Prednison wurde in für das Ausschleichen im Doppelblinddesign geeigneten Blisterpackungen geliefert.
12	Statistische Methoden	
12a	Statistische Methoden zur Bewertung der primären und sekundären Zielkriterien	<u>Primärer Wirksamkeitsendpunkt:</u> <ul style="list-style-type: none"> Der Anteil der Patienten mit anhaltender Remission zu Woche 52 wurde für die ITT-Population, die alle randomisierten Patienten umfasste, anhand des zweiseitigen Exakten Fisher-Tests mit einem Signifikanzniveau von 0,05 analysiert. Bei Erreichen statistischer Signifikanz des primären Endpunkts erfolgte die Testung der kumulativen CS-Gesamtdosis als nächstes. <u>Sekundäre Wirksamkeitsendpunkte:</u>

Item ^a	Charakteristikum	Studieninformation
		<ul style="list-style-type: none"> • Die Analyse weiterer binärer Endpunkte folgte demselben statistischen Verfahren wie die Analyse des primären Endpunkts. • Die kumulative CS-Gesamtdosis (inkl. Prednison) wurde anhand eines nichtparametrischen Wilcoxon-Rangsummentests bestimmt. • Die Time-to-Event-Analysen erfolgten anhand der Kaplan-Meier-Methode. Die Behandlungsgruppen (HR mit zugehörigem 95 %-KI und p-Wert) wurden anhand eines Cox-Proportional-Hazard-Modells verglichen. <p><u>Sicherheitsendpunkte:</u> Die Analyse der Sicherheitsendpunkte erfolgte in erster Linie deskriptiv.</p>
12b	Weitere Analysen, wie z. B. Subgruppenanalysen und adjustierte Analysen	<p><u>Subgruppenanalysen:</u> Analysen des primären Wirksamkeitsendpunkts wurden für folgende vorab definierte Subgruppen durchgeführt:</p> <ul style="list-style-type: none"> • Geschlecht (männlich, weiblich) • Abstammung (weiß, andere) • Region (1 [westliche Länder], 2 [Südamerika], 3 [Rest der Welt]) • Alter (< 70, ≥ 70 Jahre) • Gewicht zu Baseline (< 60, ≥ 60 und < 100, ≥ 100 kg) • BMI (< 25, ≥ 25 und < 30, ≥ 30 kg/m²) • CRP zu Baseline (≤ Median, > Median) • ESR zu Baseline (≤ Median, > Median) • CS-Dosis zu Baseline (≤ Median, > Median) • Anzahl der Rezidive vor Screening (1, > 1) • Von COVID-19 betroffene Population (ja, nein) – nur durchgeführt, wenn > 10 % der Patienten von COVID-19 betroffen waren
Resultate		
13	Patientenfluss (inklusive Flow-Chart zur Veranschaulichung im Anschluss an die Tabelle)	Eine Darstellung des Patientenflusses findet sich in Abbildung 4-18 im Anschluss an diese Tabelle.
13a	Anzahl der Studienteilnehmer für jede durch Randomisierung gebildete Behandlungsgruppe, die a) randomisiert wurden, b) tatsächlich die geplante Behandlung/Intervention erhalten haben, c) in der Analyse des primären Zielkriteriums berücksichtigt wurden	<p>a) Anzahl der randomisierten Patienten: N=118</p> <ul style="list-style-type: none"> • Sarilumab (+ Prednison-Ausschleichen): N=60 • Placebo (+ Prednison-Ausschleichen): N=58 <p>b) Anzahl der Patienten, die tatsächlich die geplante Behandlung erhalten haben: N=117</p> <ul style="list-style-type: none"> • Sarilumab (+ Prednison-Ausschleichen): N=59 • Placebo (+ Prednison-Ausschleichen): N=58 <p>c) Anzahl der Patienten, die in der Analyse des primären Zielkriteriums berücksichtigt wurden: N=118</p> <ul style="list-style-type: none"> • Sarilumab (+ Prednison-Ausschleichen): N=60 • Placebo (+ Prednison-Ausschleichen): N=58

Item ^a	Charakteristikum	Studieninformation
13b	Für jede Gruppe: Beschreibung von verlorenen und ausgeschlossenen Patienten nach Randomisierung mit Angabe von Gründen	<p>Verlorene oder ausgeschlossene Patienten: n=69 (58,5 %)</p> <p>Sarilumab (+ Prednison-Ausschleichen): n=32 (53,3)</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse: n=13 (21,7) % • Fehlende Wirksamkeit: n=4 (6,7 %) • Patientenentscheidung: n=3 (5,0 %) • Andere: n=12 (20,0 %) <p>Placebo (+ Prednison-Ausschleichen): n=37 (63,8 %)</p> <ul style="list-style-type: none"> • Unerwünschte Ereignisse: n=5 (8,6 %) • Fehlende Wirksamkeit: n=9 (15,5 %) • Patientenentscheidung: n=4 (6,9 %) • Andere: n=19 (32,8 %)
14	Aufnahme / Rekrutierung	
14a	Nähere Angaben über den Zeitraum der Studienaufnahme der Probanden / Patienten und der Nachbeobachtung	Die erste Visite des ersten Patienten erfolgte am 09. Oktober 2018. Die gesamte Studiendauer betrug für die einzelnen Patienten etwa 62 Wochen. Dazu zählten eine bis zu vierwöchige Screening-Phase, eine 52-wöchige doppelblinde Behandlungsphase und eine sechswöchige Nachbeobachtungsphase.
14b	Informationen, warum die Studie endete oder beendet wurde	Die Studie wurde aufgrund der Rekrutierungsproblematik während der COVID-19-Pandemie mit Datum 19. Mai 2021 (letzter Patient letzte Visite) frühzeitig beendet.
a: nach CONSORT 2010.		

Stellen Sie für jede Studie den Patientenfluss in einem Flow-Chart gemäß CONSORT dar.

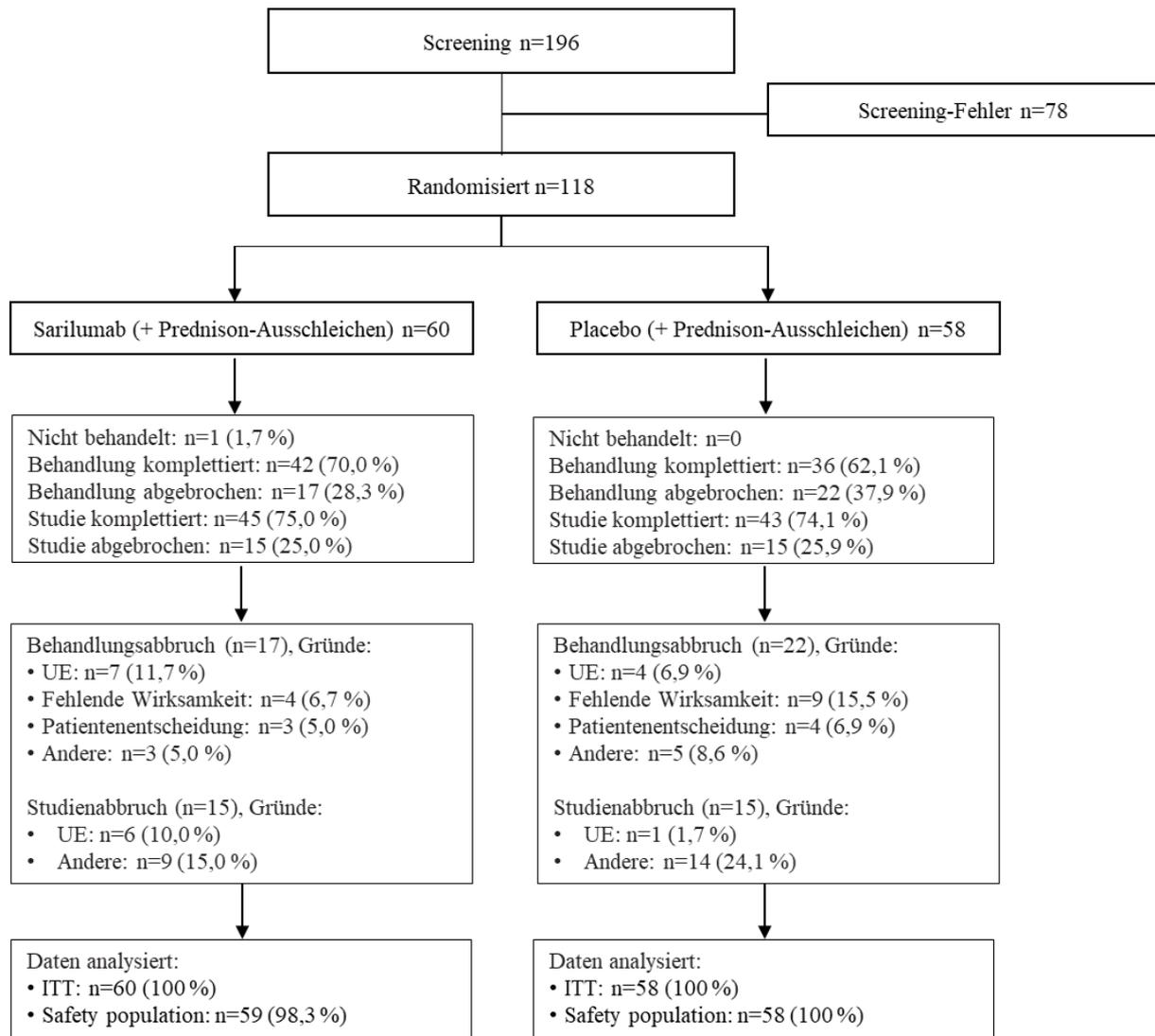


Abbildung 4-18: Flow Chart der Studie SAPHYR

Anhang 4-F: Bewertungsbögen zur Einschätzung von Verzerrungsaspekten

Der nachfolgend dargestellte Bewertungsbogen dient der Dokumentation der Einstufung des Potenzials der Ergebnisse für Verzerrungen (Bias). Für jede Studie soll aus diesem Bogen nachvollziehbar hervorgehen, inwieweit die Ergebnisse für die einzelnen Endpunkte als möglicherweise verzerrt bewertet wurden, was die Gründe für die Bewertung waren und welche Informationen aus den Quellen dafür Berücksichtigung fanden.

Der Bogen gliedert sich in zwei Teile:

- Verzerrungsaspekte auf Studienebene. In diesem Teil sind die endpunktübergreifenden Kriterien aufgelistet.
- Verzerrungsaspekte auf Endpunktebene. In diesem Teil sind die Kriterien aufgelistet, die für jeden Endpunkt separat zu prüfen sind.

Für jedes Kriterium sind unter „Angaben zum Kriterium“ alle relevanten Angaben aus den Quellen zur Bewertung einzutragen (Stichworte reichen ggf., auf sehr umfangreiche Informationen in den Quellen kann verwiesen werden).

Grundsätzlich sollen die Bögen studienbezogen ausgefüllt werden. Wenn mehrere Quellen zu einer Studie vorhanden sind, müssen die herangezogenen Quellen in der folgenden Tabelle genannt und jeweils mit Kürzeln (z. B. A, B, C ...) versehen werden. Quellenspezifische Angaben im weiteren Verlauf sind mit dem jeweiligen Kürzel zu kennzeichnen.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen ist die Blankoversion des Bogens. Dieser Blankobogen ist für jede Studie heranzuziehen. Im Anschluss daran ist ein Bewertungsbogen inklusive Ausfüllhinweisen abgebildet, der als Ausfüllhilfe dient, aber nicht als Vorlage verwendet werden soll.

Beschreiben Sie nachfolgend die Verzerrungsaspekte jeder eingeschlossenen Studie (einschließlich der Beschreibung für jeden berücksichtigten Endpunkt). Erstellen Sie hierfür je Studie eine separate Version des nachfolgend dargestellten Bewertungsbogens.

Tabelle 4-95 (Anhang): Bewertungsbogen zur Beschreibung von Verzerrungsaspekten für Studie SAPHYR

Studie: SAPHYR

Tabelle: Liste der für die Bewertung herangezogenen Quellen

Genaue Benennung der Quelle	Kürzel
Clinical Study Report: SAPHYR	B

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein → Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

Die Studie wurde randomisiert durchgeführt.

1.

für randomisierte Studien: Adäquate Erzeugung der Randomisierungssequenz

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Randomisierung erfolgte zentral mit Hilfe einer IRT.

für nicht randomisierte Studien: Zeitliche Parallelität der Gruppen

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien: Verdeckung der Gruppenzuteilung („allocation concealment“)

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Gruppenzuteilung erfolgte zufällig mit Hilfe einer IRT.

für nicht randomisierte Studien: Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient:**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Patienten waren hinsichtlich der Zuteilung der Studienmedikation (Sarilumab + CST oder Placebo + CST) verblindet. Die Prednison-Gabe erfolgte ebenfalls doppelblind.

behandelnde bzw. weiterbehandelnde Personen:

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Behandelnde bzw. weiterbehandelnde Personen waren hinsichtlich der Zuteilung der Studienmedikation (Sarilumab + CST oder Placebo + CST) verblindet. Die Prednison-Gabe erfolgte ebenfalls doppelblind.

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

In den Studienunterlagen finden sich keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen): **niedrig** **hoch**

Begründung für die Einstufung:

Randomisierung und Gruppenzuteilung erfolgten adäquat und verdeckt mittels IRT. Sowohl Patienten als auch behandelnde bzw. weiterbehandelnde Personen waren hinsichtlich Zuteilung der Studienmedikation (Sarilumab + CST oder Placebo + CST) verblindet. Es gibt keinen Anlass zur Vermutung, dass Ergebnisse selektiv berichtet wurden, da alle Analysen prospektiv im SAP definiert waren oder gemäß Verfahrensordnung des G-BA berichtet wurden. Weitere Verzerrungspotenziale liegen nicht vor.

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:**Endpunkt: Gesamtmortalität****1. Verblindung der Endpunkterheber** **ja** **unklar** **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips **ja** **unklar** **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Ergebnisse des Endpunkts Gesamtmortalität basieren auf den Daten der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten. Im gesamten Studienverlauf traten keine Todesfälle auf.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine **ja** **unklar** **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts Gesamtmortalität.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können **ja** **nein**Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts Gesamtmortalität beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig **hoch**

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt Gesamtmortalität ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten. Es traten während des gesamten Studienverlaufs keine Todesfälle auf. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Gesamtmortalität als niedrig eingestuft.

Endpunkt: Anhaltende Remission zu Woche 52

1. Verblindung der Endpunkterheber

ja **unklar** **nein**

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts anhaltende Remission zu Woche 52 basierte auf der ITT-Population, die alle randomisierten Patienten einschloss.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja **unklar** **nein**

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts anhaltende Remission zu Woche 52. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts anhaltende Remission zu Woche 52 beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt anhaltende Remission zu Woche 52 ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA durchgeführt wurden. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts anhaltende Remission zu Woche 52 als niedrig eingestuft.

Endpunkt: Kumulative Corticoid-Gesamtdosis über 52 Wochen**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts kumulative Corticoid-Gesamtdosis über 52 Wochen basierte auf der ITT-Population, die alle randomisierten Patienten einschloss und mindestens eine Prednison-Dosis erhalten hatten. Ein Patient, der keine Studienmedikation, somit auch kein Prednison, erhalten hatte, wurde von der Endpunkt-Analyse ausgeschlossen. Da dies weniger als 5 % der Patienten betraf, ist das ITT-Prinzip adäquat umgesetzt.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts kumulative Corticoid-Gesamtdosis über 52 Wochen. Alle Analysen waren prospektiv im SAP definiert.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts kumulative Corticoid-Gesamtdosis über 52 Wochen beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt kumulative Corticoid-Gesamtdosis über 52 Wochen ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss und mindestens eine Prednison-Dosis erhalten hatten. Ein Patient, der keine Studienmedikation, somit auch kein Prednison, erhalten hatte, wurde von der Endpunkt-Analyse ausgeschlossen. Da dies weniger als 5 % der Patienten betraf, ist das ITT-Prinzip umgesetzt. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts kumulative Corticoid-Gesamtdosis über 52 Wochen als niedrig eingestuft.

Endpunkt: Zeit bis zum ersten PMR-Schub

1. Verblindung der Endpunkterheber

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts Zeit bis zum ersten PMR-Schub basierte auf der ITT-Population, die alle randomisierten Patienten einschloss.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts Zeit bis zum ersten PMR-Schub. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts Zeit bis zum ersten PMR-Schub beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt Zeit bis zum ersten PMR-Schub ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA durchgeführt wurden. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Zeit bis zum ersten PMR-Schub als niedrig eingestuft.

Endpunkt: Polymyalgia rheumatica-Aktivitätsscore (PMR-AS)**1. Verblindung der Endpunkterheber**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Der PMR-AS setzt sich aus den Einzelkomponenten CRP, VAS (Patienten- und Prüfarztbeurteilung), Dauer der Morgensteifigkeit und Beweglichkeit der oberen Gliedmaßen zusammen.

Die Auswertungen der Einzelkomponenten VAS (Patienten- und Prüfarztbeurteilung) basierten für die Responderanalyse auf der ITT-Population, die alle randomisierten Patienten einschloss.

Die Auswertungen des globalen PMR-AS und der Einzelkomponenten CRP, Dauer der Morgensteifigkeit und Beweglichkeit der oberen Gliedmaßen berücksichtigt alle randomisierten Patienten, für die sowohl Baseline- als auch eine Post-Baseline-Erhebungen zu den entsprechenden Analysezeitpunkten vorlagen.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts PMR-AS. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts PMR-AS beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt PMR-AS ist als hoch einzustufen, da eine Umsetzung des ITT-Prinzips für den globalen PMR-AS und die Einzelkomponenten CRP, Dauer der Morgensteifigkeit und Beweglichkeit der oberen Gliedmaßen des Endpunkts nicht gegeben ist und somit eine Verzerrung der Ergebnisse auf Endpunktebene nicht ausgeschlossen werden kann. Es lagen keine weiteren oder sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA durchgeführt wurden. Insgesamt wird das Verzerrungspotenzial des Endpunkts PMR-AS als hoch eingestuft.

Endpunkt: Gesundheitszustand gemäß FACIT-F**1. Verblindung der Endpunkterheber**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts Gesundheitszustand gemäß FACIT-F basierte auf der ITT-Population, die alle randomisierten Patienten einschloss.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts Gesundheitszustand gemäß FACIT-F. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts Gesundheitszustand gemäß FACIT-F beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): **niedrig** **hoch**

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt Gesundheitszustand gemäß FACIT-F ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA durchgeführt wurden. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Gesundheitszustand gemäß FACIT-F als niedrig eingestuft.

Endpunkt: Gesundheitszustand gemäß EQ-5D VAS**1. Verblindung der Endpunkterheber** **ja** **unklar** **nein**Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips **ja** **unklar** **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts Gesundheitszustand gemäß EQ-5D VAS basierte auf der ITT-Population, die alle randomisierten Patienten einschloss.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine **ja** **unklar** **nein**Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts Gesundheitszustand gemäß EQ-5D VAS. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts Gesundheitszustand gemäß EQ-5D VAS beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt Gesundheitszustand gemäß EQ-5D VAS ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA durchgeführt wurden. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Gesundheitszustand gemäß EQ-5D VAS als niedrig eingestuft.

Endpunkt: Körperlicher Funktionsstatus gemäß HAQ-DI**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts körperlicher Funktionsstatus gemäß HAQ-DI basierte auf der ITT-Population, die alle randomisierten Patienten einschloss.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine ja unklar neinAngaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts körperlicher Funktionsstatus gemäß HAQ-DI. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können ja neinAngaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts körperlicher Funktionsstatus gemäß HAQ-DI beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen): niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt körperlicher Funktionsstatus gemäß HAQ-DI ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA durchgeführt wurden. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts körperlicher Funktionsstatus gemäß HAQ-DI als niedrig eingestuft.

Endpunkt: Gesundheitsbezogene Lebensqualität gemäß SF-36v2**1. Verblindung der Endpunkterheber** ja unklar neinAngaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts gesundheitsbezogene Lebensqualität gemäß SF-36v2 basierte auf der ITT-Population, die alle randomisierten Patienten einschloss.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts gesundheitsbezogene Lebensqualität gemäß SF-36v2. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts gesundheitsbezogene Lebensqualität gemäß SF-36v2 beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt gesundheitsbezogene Lebensqualität gemäß SF-36v2 ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der ITT-Population, die alle randomisierten Patienten einschloss. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA durchgeführt wurden. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts gesundheitsbezogene Lebensqualität gemäß SF-36v2 als niedrig eingestuft.

Endpunkt: Nebenwirkungen (unerwünschte Ereignisse, UE)**1. Verblindung der Endpunkterheber**

ja unklar nein

Angaben zum Kriterium; obligate Begründung für die Einstufung:

Die Studiendurchführung erfolgte doppelblind.

2. Adäquate Umsetzung des ITT-Prinzips

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Die Auswertung des Endpunkts Nebenwirkungen basierte auf der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten.

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

ja unklar nein

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

Es finden sich keine Hinweise auf eine ergebnisgesteuerte Berichterstattung des Endpunkts Nebenwirkungen. Alle Analysen waren prospektiv im SAP definiert oder wurden gemäß Verfahrensordnung des G-BA durchgeführt.

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

ja nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Es gibt keine Hinweise auf sonstige Aspekte, die das Verzerrungspotenzial des Endpunkts Nebenwirkungen beeinflussen.

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

niedrig hoch

Begründung für die Einstufung:

Das Verzerrungspotenzial für den Endpunkt Nebenwirkungen ist als niedrig einzustufen, da eine Einflussnahme der Patienten, des Arztes oder weiterer in die Behandlung involvierter Personen aufgrund des doppelblinden Studiendesigns ausgeschlossen werden kann. Die Analyse erfolgte auf Basis der Sicherheitspopulation, die alle randomisierten Patienten einschloss, die mindestens eine Dosis oder Teildosis der Studienmedikation erhalten hatten. Eine ergebnisabhängige Berichterstattung kann ebenfalls ausgeschlossen werden, da alle Analysen im SAP definiert waren oder gemäß Verfahrensordnung des G-BA durchgeführt wurden. Es lagen auch keine sonstigen Aspekte für ein hohes Verzerrungspotenzial vor. Insgesamt wird das Verzerrungspotenzial des Endpunkts Nebenwirkungen als niedrig eingestuft.

Hinweis: Der nachfolgend dargestellte Bewertungsbogen mit Ausfüllhinweisen dient nur als Ausfüllhilfe für den Blankobogen. Er soll nicht als Vorlage verwendet werden.

Bewertungsbogen zur Beschreibung von Verzerrungsaspekten (Ausfüllhilfe)

Anhand der Bewertung der folgenden Kriterien soll das Ausmaß möglicher Ergebnisverzerrungen eingeschätzt werden (A: endpunktübergreifend; B: endpunktspezifisch).

A Verzerrungsaspekte auf Studienebene:

Einstufung als randomisierte Studie

ja → Bewertung der Punkte 1 und 2 für randomisierte Studien

nein: Aus den Angaben geht klar hervor, dass es keine randomisierte Zuteilung gab, oder die Studie ist zwar als randomisiert beschrieben, es liegen jedoch Anzeichen vor, die dem widersprechen (z. B. wenn eine alternierende Zuteilung erfolgte). Eine zusammenfassende Bewertung der Verzerrungsaspekte soll für nicht randomisierte Studien nicht vorgenommen werden.

→ Bewertung der Punkte 1 und 2 für nicht randomisierte Studien

Angaben zum Kriterium:

1.

für randomisierte Studien:

Adäquate Erzeugung der Randomisierungssequenz

ja: Die Gruppenzuteilung erfolgte rein zufällig, und die Erzeugung der Zuteilungssequenz ist beschrieben und geeignet (z. B. computergenerierte Liste).

unklar: Die Studie ist zwar als randomisiert beschrieben, die Angaben zur Erzeugung der Zuteilungssequenz fehlen jedoch oder sind ungenügend genau.

nein: Die Erzeugung der Zuteilungssequenz war nicht adäquat.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:

Zeitliche Parallelität der Gruppen

ja: Die Gruppen wurden zeitlich parallel verfolgt.

unklar: Es finden sich keine oder ungenügend genaue diesbezügliche Angaben.

nein: Die Gruppen wurden nicht zeitlich parallel verfolgt.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

2.

für randomisierte Studien:**Verdeckung der Gruppenzuteilung („allocation concealment“)** **ja:** Eines der folgenden Merkmale trifft zu:

- Zuteilung durch zentrale unabhängige Einheit (z. B. per Telefon oder Computer)
- Verwendung von für die Patienten und das medizinische Personal identisch aussehenden, nummerierten oder kodierten Arzneimitteln/Arzneimittelbehältern
- Verwendung eines seriennummerierten, versiegelten und undurchsichtigen Briefumschlags, der die Gruppenzuteilung beinhaltet

 unklar: Die Angaben der Methoden zur Verdeckung der Gruppenzuteilung fehlen oder sind ungenügend genau. **nein:** Die Gruppenzuteilung erfolgte nicht verdeckt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

für nicht randomisierte Studien:**Vergleichbarkeit der Gruppen bzw. adäquate Berücksichtigung von prognostisch relevanten Faktoren** **ja:** Eines der folgenden Merkmale trifft zu:

- Es erfolgte ein Matching bzgl. der wichtigen Einflussgrößen und es gibt keine Anzeichen dafür, dass die Ergebnisse durch weitere Einflussgrößen verzerrt sind.
- Die Gruppen sind entweder im Hinblick auf wichtige Einflussgrößen vergleichbar (siehe Baseline-Charakteristika), oder bestehende größere Unterschiede sind adäquat berücksichtigt worden (z. B. durch adjustierte Auswertung oder Sensitivitätsanalyse).

 unklar: Die Angaben zur Vergleichbarkeit der Gruppen bzw. zur Berücksichtigung von Einflussgrößen fehlen oder sind ungenügend genau. **nein:** Die Vergleichbarkeit ist nicht gegeben und diese Unterschiede werden in den Auswertungen nicht adäquat berücksichtigt.Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Verblindung von Patienten und behandelnden Personen**Patient:** **ja:** Die Patienten waren verblindet. **unklar:** Es finden sich keine diesbezüglichen Angaben. **nein:** Aus den Angaben geht hervor, dass die Patienten nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

behandelnde bzw. weiterbehandelnde Personen:

ja: Das behandelnde Personal war bzgl. der Behandlung verblindet. Wenn es, beispielsweise bei chirurgischen Eingriffen, offensichtlich nicht möglich ist, die primär behandelnde Person (z. B. Chirurg) zu verblinden, wird hier beurteilt, ob eine angemessene Verblindung der weiteren an der Behandlung beteiligten Personen (z. B. Pflegekräfte) stattgefunden hat.

unklar: Es finden sich keine diesbezüglichen Angaben.

nein: Aus den Angaben geht hervor, dass die behandelnden Personen nicht verblindet waren.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

4. Ergebnisunabhängige Berichterstattung aller relevanten Endpunkte

Falls die Darstellung des Ergebnisses eines Endpunkts von seiner Ausprägung (d. h. vom Resultat) abhängt, können erhebliche Verzerrungen auftreten. Je nach Ergebnis kann die Darstellung unterlassen worden sein (a), mehr oder weniger detailliert (b) oder auch in einer von der Planung abweichenden Weise erfolgt sein (c).

Beispiele zu a und b:

- *Der in der Fallzahlplanung genannte primäre Endpunkt ist nicht / unzureichend im Ergebnisteil aufgeführt.*
- *Es werden (signifikante) Ergebnisse von vorab nicht definierten Endpunkten berichtet.*
- *Nur statistisch signifikante Ergebnisse werden mit Schätzern und Konfidenzintervallen dargestellt.*
- *Lediglich einzelne Items eines im Methodenteil genannten Scores werden berichtet.*

Beispiele zu c: Ergebnisgesteuerte Auswahl in der Auswertung verwendeter

- *Subgruppen*
- *Zeitpunkte/-räume*
- *Operationalisierungen von Zielkriterien (z. B. Wert zum Studienende anstelle der Veränderung zum Baseline-Wert; Kategorisierung anstelle Verwendung stetiger Werte)*
- *Distanzmaße (z. B. Odds Ratio anstelle der Risikodifferenz)*
- *Cut-off-points bei Dichotomisierung*
- *statistischer Verfahren*

Zur Einschätzung einer potenziell vorhandenen ergebnisgesteuerten Berichterstattung sollten folgende Punkte – sofern möglich – berücksichtigt werden:

- *Abgleich der Angaben der Quellen zur Studie (Studienprotokoll, Studienbericht, Registerbericht, Publikationen).*
- *Abgleich der Angaben im Methodenteil mit denen im Ergebnisteil. Insbesondere eine stark von der Fallzahlplanung abweichende tatsächliche Fallzahl ohne plausible und ergebnisunabhängige Begründung deutet auf eine selektive Beendigung der Studie hin.*

Zulässige Gründe sind:

- *erkennbar nicht ergebnisgesteuert, z. B. zu langsame Patientenrekrutierung*
- *Fallzahladjustierung aufgrund einer verblindeten Zwischenauswertung anhand der Streuung der Stichprobe*
- *geplante Interimanalysen, die zu einem vorzeitigen Studienabbruch geführt haben*
- *Prüfen, ob statistisch nicht signifikante Ergebnisse weniger ausführlich dargestellt sind.*
- *Ggf. prüfen, ob „übliche“ Endpunkte nicht berichtet sind.*

Anzumerken ist, dass Anzeichen für eine ergebnisgesteuerte Darstellung eines Endpunkts zu Verzerrungen der Ergebnisse der übrigen Endpunkte führen kann, da dort ggf. auch mit einer selektiven Darstellung gerechnet werden muss. Insbesondere bei Anzeichen dafür, dass die Ergebnisse einzelner Endpunkte selektiv nicht berichtet werden, sind Verzerrungen für die anderen Endpunkte möglich. Eine von der Planung abweichende selektive Darstellung des Ergebnisses eines Endpunkts führt jedoch nicht zwangsläufig zu einer Verzerrung der anderen Endpunkte; in diesem Fall ist die ergebnisgesteuerte Berichterstattung endpunktspezifisch unter Punkt B.3 (siehe unten) einzutragen. Des Weiteren ist anzumerken, dass die Berichterstattung von unerwünschten Ereignissen üblicherweise ergebnisabhängig erfolgt (es werden nur Häufungen / Auffälligkeiten berichtet) und dies nicht zur Verzerrung anderer Endpunkte führt.

- ja:** Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.
- unklar:** Die verfügbaren Angaben lassen eine Einschätzung nicht zu.
- nein:** Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor, die das Verzerrungspotenzial aller relevanten Endpunkte beeinflusst.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

5. Keine sonstigen (endpunktübergreifenden) Aspekte, die zu Verzerrung führen können

z. B.

- zwischen den Gruppen unterschiedliche Begleitbehandlungen außerhalb der zu evaluierenden Strategien
- intransparenter Patientenfluss
- Falls geplante Interimanalysen durchgeführt wurden, so sind folgende Punkte zu beachten:
 - Die Methodik muss exakt beschrieben sein (z. B. alpha spending approach nach O'Brien Fleming, maximale Stichprobengröße, geplante Anzahl und Zeitpunkte der Interimanalysen).
 - Die Resultate (p-Wert, Punkt- und Intervallschätzung) des Endpunktes, dessentwegen die Studie abgebrochen wurde, sollten adjustiert worden sein.
 - Eine Adjustierung sollte auch dann erfolgen, wenn die maximale Fallzahl erreicht wurde.
 - Sind weitere Endpunkte korreliert mit dem Endpunkt, dessentwegen die Studie abgebrochen wurde, so sollten diese ebenfalls adäquat adjustiert werden.

- ja**
- nein**

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse auf Studienebene (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials der Ergebnisse erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen Punkte A.1 bis A.5. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

- niedrig:** Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse durch diese endpunktübergreifenden Aspekte relevant verzerrt sind.

- hoch:** Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

B Verzerrungsaspekte auf Endpunktebene pro Endpunkt:

Die folgenden Punkte B.1 bis B.4 dienen der Einschätzung der endpunktspezifischen Aspekte für das Ausmaß möglicher Ergebnisverzerrungen. Diese Punkte sollten i. d. R. für jeden relevanten Endpunkt separat eingeschätzt werden (ggf. lassen sich mehrere Endpunkte gemeinsam bewerten, z. B. Endpunkte zu unerwünschten Ereignissen).

Endpunkt: _____

1. Verblindung der Endpunkterheber

Für den Endpunkt ist zu bestimmen, ob das Personal, welches die Zielkriterien erhoben hat, bzgl. der Behandlung verblindet war.

In manchen Fällen kann eine Verblindung auch gegenüber den Ergebnissen zu anderen Endpunkten (z. B. typischen unerwünschten Ereignissen) gefordert werden, wenn die Kenntnis dieser Ergebnisse Hinweise auf die verabreichte Therapie gibt und damit zu einer Entblindung führen kann.

- ja:** Der Endpunkt wurde verblindet erhoben.
- unklar:** Es finden sich keine diesbezüglichen Angaben.
- nein:** Aus den Angaben geht hervor, dass keine verblindete Erhebung erfolgte.

Angaben zum Kriterium; obligate Begründung für die Einstufung:

2. Adäquate Umsetzung des ITT-Prinzips

Kommen in einer Studie Patienten vor, die die Studie entweder vorzeitig abgebrochen haben oder wegen Protokollverletzung ganz oder teilweise aus der Analyse ausgeschlossen wurden, so sind diese ausreichend genau zu beschreiben (Abbruchgründe, Häufigkeit und Patientencharakteristika pro Gruppe) oder in der statistischen Auswertung angemessen zu berücksichtigen (i. d. R. ITT-Analyse, siehe Äquivalenzstudien). Bei einer ITT („intention to treat“)-Analyse werden alle randomisierten Patienten entsprechend ihrer Gruppenzuteilung ausgewertet (ggf. müssen fehlende Werte für die Zielkriterien in geeigneter Weise ersetzt werden). Zu beachten ist, dass in Publikationen der Begriff ITT nicht immer in diesem strengen Sinne Verwendung findet. Es werden häufig nur die randomisierten Patienten ausgewertet, die die Therapie zumindest begonnen haben und für die mindestens ein Post-Baseline-Wert erhoben worden ist („full analysis set“). Dieses Vorgehen ist in begründeten Fällen Guideline-konform, eine mögliche Verzerrung sollte jedoch, insbesondere in nicht verblindeten Studien, überprüft werden. Bei Äquivalenz- und Nichtunterlegenheitsstudien ist es besonders wichtig, dass solche Patienten sehr genau beschrieben werden und die Methode zur Berücksichtigung dieser Patienten transparent dargestellt wird.

- ja:** Eines der folgenden Merkmale trifft zu:
- Laut Studienunterlagen sind keine Protokollverletzer und Lost-to-follow-up-Patienten in relevanter Anzahl (z. B. Nichtberücksichtigungsanteil in der Auswertung < 5 %) aufgetreten, und es gibt keine Hinweise (z. B. diskrepante Patientenzahlen in Flussdiagramm und Ergebnistabelle), die dies bezweifeln lassen.

- Die Protokollverletzer und Lost-to-follow-up-Patienten sind so genau beschrieben (Art, Häufigkeit und Charakteristika pro Gruppe), dass deren möglicher Einfluss auf die Ergebnisse abschätzbar ist (eigenständige Analyse möglich).
- Die Strategie zur Berücksichtigung von Protokollverletzern und Lost-to-follow-up-Patienten (u. a. Ersetzen von fehlenden Werten, Wahl der Zielkriterien, statistische Verfahren) ist sinnvoll angelegt worden (verzerrt die Effekte nicht zugunsten der zu evaluierenden Behandlung).

unklar: Aufgrund unzureichender Darstellung ist der adäquate Umgang mit Protokollverletzern und Lost-to-follow-up-Patienten nicht einschätzbar.

nein: Keines der unter „ja“ genannten drei Merkmale trifft zu.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

3. Ergebnisunabhängige Berichterstattung dieses Endpunkts alleine

Beachte die Hinweise zu Punkt A.4!

ja: Eine ergebnisgesteuerte Berichterstattung ist unwahrscheinlich.

unklar: Die verfügbaren Angaben lassen eine Einschätzung nicht zu.

nein: Es liegen Anzeichen für eine ergebnisgesteuerte Berichterstattung vor.

Angaben zum Kriterium; falls unklar oder nein, obligate Begründung für die Einstufung:

4. Keine sonstigen (endpunktspezifischen) Aspekte, die zu Verzerrungen führen können

z. B.

- *relevante Dateninkonsistenzen innerhalb der oder zwischen Studienunterlagen*
- *unplausible Angaben*
- *Anwendung inadäquater statistischer Verfahren*

ja

nein

Angaben zum Kriterium; falls nein, obligate Begründung für die Einstufung:

Einstufung des Verzerrungspotenzials der Ergebnisse des Endpunkts (ausschließlich für randomisierte Studien durchzuführen):

Die Einstufung des Verzerrungspotenzials erfolgt unter Berücksichtigung der einzelnen Bewertungen der vorangegangenen endpunktspezifischen Punkte B.1 bis B.4 sowie der Einstufung des Verzerrungspotenzials auf Studienebene. Falls die endpunktübergreifende Einstufung mit „hoch“ erfolgte, ist das Verzerrungspotenzial für den Endpunkt i. d. R. auch mit „hoch“ einzuschätzen. Eine relevante Verzerrung bedeutet hier, dass sich die Ergebnisse bei Behebung der verzerrenden Aspekte in ihrer Grundaussage verändern würden.

niedrig: Es kann mit großer Wahrscheinlichkeit ausgeschlossen werden, dass die Ergebnisse für diesen Endpunkt durch die endpunktspezifischen sowie endpunktübergreifenden Aspekte relevant verzerrt sind.

hoch: Die Ergebnisse sind möglicherweise relevant verzerrt.

Begründung für die Einstufung:

Anhang 4-G: Weitere Analysen

Siehe separates Dokument.